

OPTICAL NETWORKS
Biswanath Mukherjee *Series Editor*

Rudra Dutta
Ahmed E. Kamal
George N. Rouskas *Editors*

Traffic Grooming for Optical Networks

Foundations, Techniques,
and Frontiers

 Springer

Traffic Grooming for Optical Networks

Optical Networks

Series Editor: Biswanath Mukherjee
University of California, Davis
Davis, CA

Traffic Grooming for Optical Networks: Foundations, Techniques, and Frontiers
Rudra Dutta, Ahmed E. Kamal, and George N. Rouskas (Eds.)
ISBN 978-0-387-74517-6

Optical Network Design and Planning
Jane M. Simmons
ISBN 978-0-387-76475-7

Quality of Service in Optical Burst Switched Networks
Kee Chaing Chua, Mohan Gurusamy, Yong Liu, and Minh Hoang Phung
ISBN 978-0-387-34160-6

Optical WDM Networks
Biswanath Mukherjee
ISBN 978-0-387-29055-3

Traffic Grooming in Optical WDM Mesh Networks
Keyao Zhu, Hongyue Zhu, and Biswanath Mukherjee
ISBN 978-0-387-25432-6

Survivable Optical WDM Networks
Canhui (Sam) Ou and Biswanath Mukherjee
ISBN 978-0-387-24498-3

Optical Burst Switched Networks
Jason P. Jue and Vinod M. Vokkarane
ISBN 978-0-387-23756-5

Rudra Dutta · Ahmed E. Kamal ·
George N. Rouskas
Editors

Traffic Grooming for Optical Networks

Foundations, Techniques, and Frontiers

 Springer

Editors

Rudra Dutta
North Carolina State University
Department of Computer Science
Raleigh, NC
USA

Ahmed E. Kamal
Iowa State University
Department of Electrical
and Computer Engineering
Ames, IA
USA

George N. Rouskas
North Carolina State University
Department of Computer Science
Raleigh, NC
USA

ISSN: 1935-3839 e-ISSN: 1935-3847
ISBN: 978-0-387-74517-6 e-ISBN: 978-0-387-74518-3
DOI: 10.1007/978-0-387-74517-6

Library of Congress Control Number: 2008938757

© 2008 Springer Science+Business Media, LLC

All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Springer Science+Business Media, LLC, 233 Spring Street, New York, NY 10013, USA), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden.

The use in this publication of trade names, trademarks, service marks, and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

Printed on acid-free paper

springer.com

To Sambhu and Ira, my parents,
To Amrapali, my wife,
To Mouli, my sister,
– all my best teachers
RD

To my wife, Nermin,
and our children, Mohammed, Mona, Heba and Youssef
AEK

To my wonderful wife, Magdalini
GNR

Preface

The objective of this book is to provide timely and comprehensive coverage of the principles, technology, practice, and future of traffic grooming in optical networks. Traffic grooming considerations are already shaping new switch designs and standards, including next generation Synchronous Optical Network/Synchronous Digital Hierarchy (SONET/SDH) and Generic Framing Procedure (GFP), are affecting future optical network technologies, and are creating new business opportunities. Yet information on the topic is scattered and there is a shortage of technical sources where relevant material can be accessed in a single location. The motivation for this book was to bridge this gap by providing a single authoritative point of reference to the traffic grooming state of the art.

Traffic grooming is a complex subject, involving a number of interrelated concepts, standards, and technologies. It is also a rapidly growing field of study, making it difficult for a single book to cover all aspects in detail. To cope with this scope and complexity, this book provides appropriate background information, followed by an in-depth study of a few key issues and challenges. It has been our intention to bring together a broad range of perspectives from preeminent researchers in both academia and industry. We believe that these multiple, diverse points of view add considerable value and make the contents more interesting to the reader.

This book is intended for practicing engineers as well as industry and academic researchers. The potential audience includes new entrants to the field, including industry practitioners and graduate students in computer science, telecommunications, and related disciplines interested in practical information on traffic grooming; network designers and planners and engineering managers involved in cross-connect design, optical network design, and standardization efforts who are interested in traffic grooming technologies and techniques; and researchers who wish to explore the subject matter further. The book is also suitable as textbook for graduate-level courses on optical networks or network design, as well as for industry short courses on traffic grooming, and as a comprehensive reference for those conducting original research in the field.

Book Organization

The Introduction provides an easy entry to the field of traffic grooming, even for readers unfamiliar with optical networking. The remainder of the book is divided into three parts. The first part provides essential background material and the corresponding chapters are more tutorial in nature, whereas the last two parts examine several key research issues in depth.

Part I, *Foundations*, presents enabling technology and standards, and provides a formal introduction to traffic grooming theory. It contains five chapters covering grooming switch architectures, related control plane standards, the grooming capabilities of SONET and next-generation SONET technologies, the computational complexity of the problem, and a scalable hierarchical grooming framework.

Part II, *Techniques*, consists of eight chapters which present specific traffic grooming techniques addressing corresponding fundamental issues including grooming in SONET and next generation SONET; survivability; mathematical programming approaches for static grooming; grooming under scheduled service and dynamic traffic; performance modeling; and multipoint grooming.

Finally, Part III, *Frontiers*, examines topics of emerging importance, including multidomain considerations, waveband grooming and switching, and all-optical grooming.

Acknowledgments

We are grateful to the contributing authors for their efforts and diligence. Naturally, this book would not have been possible without their help and cooperation – our heartfelt thanks to them.

We also acknowledge the National Science Foundation which has provided partial support for our work on this project under grants CNS-0322107 and CNS-0626741.

We thank Springer Science+Business Media for the opportunity to publish this book. We are especially grateful to Professor Biswanath Mukherjee for his encouragement and support during the process of writing and editing this book. We also thank Alex Greene and Katie Stanne at Springer Science+Business Media for their patience and help in guiding us through the maze of this task.

Raleigh, NC, USA
Ames, IA, USA
Releigh, NC, USA

Rudra Dutta
Ahmed E. Kamal
George N. Rouskas

Contents

1 Introduction

Rudra Dutta, Ahmed E. Kamal, George N. Rouskas 1

Part I Foundations

2 Grooming Switches

Tarek S. El-Bawab 9

3 Control Plane Support

Slobodanka Tomic 21

4 Grooming Mechanisms in SONET/SDH and Next-Generation SONET/SDH

Rudra Dutta, Ahmed E. Kamal, George N. Rouskas 39

5 Computational Complexity

Shu Huang 57

6 Hierarchical Traffic Grooming

George N. Rouskas, Rudra Dutta 73

Part II Techniques

7 Traffic Grooming in SONET/SDH Rings

Raza Ul-Mustafa 91

8 Traffic Grooming in Next-Generation SONET/SDH Networks

Smita Rai, Biswanath Mukherjee 107

9 Mathematical Programming Approaches
JianQiang Hu 123

10 Survivable Traffic Grooming
Arun K. Somani 137

11 Traffic Grooming Under Scheduled Service
Bin Wang..... 159

12 Dynamic Grooming Algorithms
Byrav Ramamurthy 175

13 Performance Models for Dynamic Traffic Grooming
Chunsheng Xin 193

14 Multipoint Traffic Grooming
Ahmed E. Kamal, Ahmad N. Khalil 213

Part III Frontiers

15 Multi-Domain Traffic Grooming
Nasir Ghani, Qing Liu, Nageswara S. V. Rao, Ashwin Gumaste 237

16 Grooming of Scheduled Demands in Multi-Layer Optical Networks
Maurice Gagnaire, Josué Kuri, Elias A. Doumith 253

17 All-Optical Traffic Grooming
Volkan Kaman, Henrik N. Poulsen, Roger J. Helkey, John E. Bowers... 279

Index 291

List of Contributors

John E. Bowers
Calient Networks
25 Castilian Drive
Goleta, CA 93117, USA
jbowers@calient.net

Elias A. Doumith
Department of Computer Science
and Networks
École Nationale Supérieure des
Télécommunications
46, rue Barrault, 75013 Paris, France
elias.doumith@enst.fr

Rudra Dutta
Department of Computer Science
North Carolina State University
890 Oval Drive
Raleigh, NC 27695-8206, USA
dutta@csc.ncsu.edu

Tarek El-Bawab
Department of Computer Engg.
Jackson State University
Jackson, MS 39217, USA
tarek.el-bawab@jsums.edu

Maurice Gagnaire
Department of Computer Science
and Networks
École Nationale Supérieure des

Télécommunications
46, rue Barrault, 75013 Paris, France
maurice.gagnaire@enst.fr

Nasir Ghani
Department of Electrical and
Computer Engineering
University of New Mexico
Albuquerque, NM 87131, USA
nghani@ece.unm.edu

Ashwin Gumaste
Department of Computer Science
and Engineering
Indian Institute of Technology
Bombay
Powai, Mumbai, India 400076
ashwing@ieee.org

Roger J. Helkey
Calient Networks
25 Castilian Drive
Goleta, CA 93117, USA
rhelkey@calient.net

JianQiang Hu
Department of Manufacturing
Engineering
Boston University
15 St. Mary's Street
Brookline, MA 02446, USA
hqliang@bu.edu

Shu Huang

Renaissance Computing Institute
100 Europa Drive Suite 540
Chapel Hill, NC 27517, USA
shuang@renci.org

Ahmed E. Kamal

Department of Electrical and
Computer Engineering
Iowa State University
Ames, IA 50011, USA
kamal@iastate.edu

Volkan Kaman

Calient Networks
25 Castilian Drive
Goleta, CA 93117, USA
vkaman@calient.net

Ahmad N. Khalil

City University of New York
New York, NY 10031, USA
akhilil@ccny.cuny.edu

Josué Kuri

Orange Labs
San Francisco, USA
josue.kuri@orange-ftgroup.com

Qing Liu

Department of Electrical and
Computer Engineering
University of New Mexico
Albuquerque, NM 87131, USA

Biswanath Mukherjee

Department of Computer Science
University of California
Davis, CA 95616, USA
mukherje@cs.ucdavis.edu

Henrik N. Poulsen

Department of Electrical and
Computer Engineering
University of California, Santa
Barbara
Santa Barbara, CA 93106, USA
henrik@ece.ucsb.edu

Smita Rai

Department of Computer Science
University of California
Davis, CA 95616, USA
rai@cs.ucdavis.edu

Byrav Ramamurthy

Department of Computer Science
and Engineering
University of Nebraska-Lincoln
Lincoln, NE 68588, USA
byrav@cse.unl.edu

Nageswara S. V. Rao

Oak Ridge National Laboratory
PO Box 2008 MS6015
Oak Ridge TN 37831, USA
raons@ornl.gov

George N. Rouskas

Department of Computer Science
North Carolina State University
890 Oval Drive
Raleigh, NC 27695-8206, USA
rouskas@ncsu.edu

Arun K. Somani

Department of Electrical and
Computer Engineering
Iowa State University
Ames, IA 50011, USA
arun@iastate.edu

Slobodanka Tomic

Telecommunications Center Vienna
(ftw.)
Donau-City-Strasse 1
A-1220 Vienna, Austria
tomic@ftw.at

Raza Ul-Mustafa

Microsoft Corporation
One Microsoft Way
Redmond, WA 98052, USA
razaulm@microsoft.com

Bin Wang

Department of Computer Science
Wright State
University Dayton, OH 45435, USA
bin.wang@wright.edu

Chunsheng Xin

Department of Computer Science
Norfolk
State University 700 Park Ave.
Norfolk, VA 23504, USA
cxin@nsu.edu

Introduction

Rudra Dutta, Ahmed E. Kamal, George N. Rouskas

Computer networking has certainly changed the face of business, education, and society in general in the last half a century. The increasing impact of computer networks can be traced to the twin advances of *where* the network can provide service and *how much* the network can do. The first advance was started by the initial deployment of the Internet in its various incarnations and carried on, more recently, by the advent of ubiquitous networking, utilizing wireless technology in various forms. The second is an equally dramatic change, in terms of not only the volume of bits that can be transferred, but also the predictability with which this can be accomplished. The concepts of QoS and SLAs have emerged as the new metrics to measure the usefulness of the network, in addition to more fundamental metrics such as throughput. In providing very high performance in both throughput and QoS, optical networking technology has come to the fore in the last two decades.

The development of communication technology using fiber-optic links and lasers for signal transmission increased the quality of data communication dramatically on several fronts. The effect was to make much higher bitrates achievable than before. However, the term “networking” is usually reserved for algorithms and protocols that enable connectivity over a large set of stations which are not all connected directly to each other; that is, operating at the third layer of the OSI model and performing forwarding and routing. The term “optical networking”, therefore, is sometimes a little confusing, since it might appear that the impact of optical transmission technology is limited to the physical layer only. However, for several reasons that have become apparent over the years, considerations of physical layer characteristics cannot be ignored in network layer design. The core of such considerations is the fact that the hallowed layering principle, although a very powerful and useful one, serves best when viewed as general guidelines rather than inflexible shackles. Like all good things, layer abstraction can be overdone.

This has become more generally recognized and acknowledged in recent years, so much so that “cross-layer” has become a generally understood term

Introduction

Rudra Dutta, Ahmed E. Kamal, George N. Rouskas

Computer networking has certainly changed the face of business, education, and society in general in the last half a century. The increasing impact of computer networks can be traced to the twin advances of *where* the network can provide service and *how much* the network can do. The first advance was started by the initial deployment of the Internet in its various incarnations and carried on, more recently, by the advent of ubiquitous networking, utilizing wireless technology in various forms. The second is an equally dramatic change, in terms of not only the volume of bits that can be transferred, but also the predictability with which this can be accomplished. The concepts of QoS and SLAs have emerged as the new metrics to measure the usefulness of the network, in addition to more fundamental metrics such as throughput. In providing very high performance in both throughput and QoS, optical networking technology has come to the fore in the last two decades.

The development of communication technology using fiber-optic links and lasers for signal transmission increased the quality of data communication dramatically on several fronts. The effect was to make much higher bitrates achievable than before. However, the term “networking” is usually reserved for algorithms and protocols that enable connectivity over a large set of stations which are not all connected directly to each other; that is, operating at the third layer of the OSI model and performing forwarding and routing. The term “optical networking”, therefore, is sometimes a little confusing, since it might appear that the impact of optical transmission technology is limited to the physical layer only. However, for several reasons that have become apparent over the years, considerations of physical layer characteristics cannot be ignored in network layer design. The core of such considerations is the fact that the hallowed layering principle, although a very powerful and useful one, serves best when viewed as general guidelines rather than inflexible shackles. Like all good things, layer abstraction can be overdone.

This has become more generally recognized and acknowledged in recent years, so much so that “cross-layer” has become a generally understood term

in network design. Today, nobody would suggest designing a network layer or a transport protocol for use in a multihop wireless network without taking wireless characteristics into account. Similarly, the optical physical layer forces itself into many networking concerns. Examples include the design of logical topologies with constraints imposed by impediments specific to optical transmission systems, the design of link layer protocols and architectures with the knowledge that the ratio of transmission delay to propagation delay is dramatically smaller than in previous wired networks, and the design of network survivability with the disparity in speed and granularity of optical layer versus higher layer restoration. Such issues, recognized now as cross-layer design issues, have been pursued by the optical networking research and development community for years. Another excellent example of such “typically optical” networking research area is traffic grooming, the topic of this book.

Traffic grooming first came to be recognized as a research area in the mid-1990s; possibly the first use of the term “grooming” in this connection was by Sasaki et al. in 1998, but earlier work by the same researchers as well as others can be retrospectively seen to address the same general design problem. While many variations of the problem have been investigated and many different flavors have been propounded, it is possible to recognize a core problem, in the light of which all others can be discussed. This core problem is essentially a network design problem of resource allocation.

In the essential grooming problem, the resource to be allocated is the switching capacity at each node of a network. The network is assumed to be composed of physical links of optical fiber that are already in place, and the demand on the network is in the form of a static traffic demand matrix that specifies a required rate of flow of traffic for each pair of nodes in the network. Naturally, to operate the network to satisfy the traffic demands, a routing of the traffic flows onto the physical topology of fibers must be obtained. However, the grooming problem is a multi-level routing problem, and one in which the goal reflects an emergent cost function rather than a direct one, as we explain below.

With Wavelength Division Multiplexing (WDM), it is possible to multiplex several optical channels of different wavelengths onto the same fiber. Optical switching technology allows an optical switch, otherwise known as an Optical Cross-Connect (OXC), to forward the optical signal arriving on an incoming fiber at the physical layer, without the use of digital electronics, onto an outgoing fiber. Further, such optical switching can be performed in a wavelength selective manner; two optical channels from the same incoming fiber can be routed, at the physical layer, onto different outgoing fiber links. Such optical switching over a series of switches forms *lightpaths* or clear optical channels. The collection of such lightpaths formed in the network is called a *virtual topology*, otherwise known as *logical topology* (because the lightpaths are *logical links*, as opposed to the physical fiber links). Specifying the sequence of physical fibers that each lightpath traverses, and specifying a

specific wavelength channel that it will use on these fibers, is called Routing and Wavelength Assignment (RWA) of the virtual topology. Unless nodes are equipped with wavelength converters, optical signals can only be switched on the same wavelength channel; this is known as the wavelength continuity constraint on lightpaths.

To inject traffic into the lightpath at the source node and to extract it from the lightpath at the destination node require the use of digital logic, and thus the nodes must be equipped with Digital Cross-Connects (DXC). With DXCs, it is also possible to extract traffic from a lightpath, and instead of forwarding it, inject it into another lightpath. Thus a second, electronic, level of traffic forwarding is possible. Such forwarding is referred to as Opto-Electro-Optic (OEO) switching, tracing the traffic through the switching node. OEO switching should not be needed if traffic between any pair of nodes is routed on its own lightpath. However, there is very likely to be a rate mismatch between the traffic demands and the lightpath capacity that provided one of the initial drivers for traffic grooming.

With currently available optical technology, the data rate of each wavelength is in the order of 2.5–10 Gbps, while channels operating at 40 Gbps and beyond will be commercially available in the near future. These rates are large for single source-to-destination channels even for backbone networks. To utilize efficiently this capacity, a number of independent lower-rate traffic streams must be multiplexed into a single lightpath. In that case, all traffic components cannot be carried on a single lightpath from source to destination, and some electronic routing of traffic with OEO conversion becomes inevitable. However, a backbone node utilizing OEO introduces delay, variability of delay, loss of throughput, and increased probability of errors, thus there is good reason to keep electronic routing to a minimum. Moreover, with the deployment of commercial WDM systems, it has become apparent that the cost of the DXCs used for OEO switching can be one of the dominant costs in building optical networks. The amount of traffic routed electronically, and thus the above direct and indirect costs to the network, will depend on the exact virtual topology formed and the mapping of lower-rate traffic components onto the lightpaths. Hence the concept of *grooming* traffic into available wavelengths arises to meet network cost or other design goals.

If the optical network is viewed as a directed graph $G_p = (V, E_p)$ where each edge represents an optical fiber link between its endpoints, then the traffic grooming problem, given a number of wavelengths W supported by each fiber and a grooming factor C , asks for the solution which optimizes some cost metric (such as the total DXC capacity over all nodes) to enable the routing of a given static traffic demand matrix $T = [t_{ij}]$. The t_{ij} are integer values expressing the aggregate traffic from node i to j in multiples of some base rate, and C is the bandwidth of a single wavelength in multiples of the same unit. Thus source-to-destination traffic components are in general *sub wavelength* in nature. The solution consists of:

1. *Specifying the lightpath set*, a set of lightpath requests, denoted by G_v (V, E_v); the virtual connection graph on the same set of nodes as the physical topology,
2. *Solving the RWA problem*, providing the routing \mathcal{R} mapping each lightpath in G_v to a sequence of links of G_p forming a path from the source of that lightpath to its destination, and the wavelength assignment \mathcal{L} , assigning a wavelength to each lightpath out of a set of W wavelengths, such that no directed link is traversed by more than W lightpaths, no directed link is traversed by more than one lightpath of every given wavelength, and (optionally) the wavelength continuity constraint is obeyed,
3. *Routing or grooming traffic*, specifying a grooming solution \mathcal{G} mapping each element of T to a sequence of lightpaths in G_v from the source of that traffic component to its destination, such that no lightpath is assigned total traffic more than C ,

such that the requirement on the cost metric is met. We represent this pictorially in Fig. 1.1. Thus the input to the problem is T and G_p , and the problem is to map T onto G_p . The solution requires specifying \mathcal{R} , \mathcal{L} , G_v , and \mathcal{G} , which complete the mapping.

Many variations of the original grooming problems have been studied in the literature. Early work concentrated on static versions of the problem, such as we have presented above, where all traffic demands are known to be reasonably unvarying over time, and embodied in a traffic matrix. In most of these variations, the problems are known to be or conjectured to be NP-hard. The literature also includes many heuristic algorithms for traffic grooming.

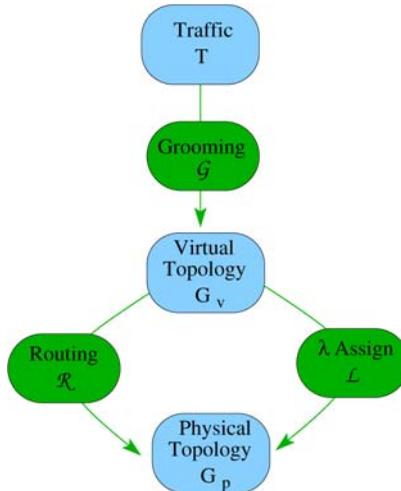


Fig. 1.1. A schematic representation of the grooming problem

Research has not been confined entirely to better algorithm design for the original problem. As the field of research has matured and responded to the changing needs as seen and foreseen by the practitioner community, the concept of traffic grooming has been carried to different arenas, and literature in this area has diverged. Comparatively recently, lower level networks in which a dynamic traffic model is more realistic have gained attention from the grooming community. It is important to note that the focus of grooming traffic shifts as a consequence of the above change in capabilities and design strategy. Reduction of OEO interchanging costs may continue to be an objective of traffic grooming. But the primary objective may now well be a minimization of the blocking behavior of the network; this is not particularly relevant in static traffic grooming because, with good planning, the entire traffic matrix is expected to be carried by the network. But making a similar 100% guarantee under statistically described dynamic traffic may be prohibitive in cost and not desirable. Similarly, the consideration of fairness is not relevant for the static problem, but may become an important one for the dynamic case.

For dynamic traffic, the grooming problem must be seen as one of supplying a *policy design* for the network, that is, an algorithm that the network control plane can employ to make decisions in response to traffic change events. The state space consists of the current virtual topology, RWA solution, and subwavelength traffic routing. The action space ranges from admission control to network layer routing, virtual topology modification, and subwavelength traffic component rearrangement.

Several traffic variation models are worth considering in dynamic traffic grooming, and the adoption of different models has marked the further evolution of the field, as researchers sought to understand and reflect real-world grooming scenarios. Subwavelength traffic components can be viewed simply as calls, using an arrival/departure model, Poisson or otherwise. It is also possible to view end-to-end traffic demands which are long-lived and do not completely depart, but undergo changes in magnitude from time to time, giving rise to an increment/decrement variation model. Another model that has been promising is the scheduled window model, in which traffic demands arrive or are scheduled, but provide a window within which they must be served, rather than immediately or as soon as possible. In an echo of the busy hour model of yesteryear's telephony networks, it is possible to model variation by providing entire traffic matrices. Network traffic demand is always one of these matrices, but changes from one to the other at unpredictable or statistically predictable times. The times may be completely predictable, in which case all traffic can be seen as predefined as scheduled, and the problem can in fact be modeled as a static problem.

The connection with the work in the protocol and signaling community is worth remarking upon. MPLS/GMPLS or ASON frameworks provide mechanisms which can be used to perform static or dynamic grooming. However, these developments have focused (as appropriate for the role of protocol standardization bodies) on enabling technology rather than design strategies. The

network administrator is provided with mechanisms to set up TE or QoS actions; but what actions are to be taken is left up to the administrator, who must look elsewhere for algorithms that provide policy or strategy decisions. In this sense, research such as traffic grooming provides a necessary complement to the development of enabling technology; in turn, the grooming researcher must understand what tools are realistically available now, or likely to be in the near future.

In this book, we have attempted a gathering together of many of these threads that connect to the central concern of network design and resource allocation with both electronic and optical switching, and typically sub-wavelength demands. We cannot claim that we have covered the entire ground in such a divergent field, but we hope that we have come very close to it; the final judgment is up to the reader. At the risk of a little overlap, we have allowed each chapter to articulate its context within this backdrop, so that each may be read by itself without too much reference back and forth. This flow of individual topics along the same channel before diverging to their separate goals seemed appropriate for the field of traffic grooming.

Foundations

Grooming Switches

Tarek S. El-Bawab

2.1 Introduction

The rising popularity of the Internet, Voice over Internet Protocol (VoIP), IP Television (IP TV), and numerous packet-based services have changed the landscape of the telecommunications industry. Several new paradigms are evolving in terms of technology, applications, services, and business models. Meanwhile, most carrier core networks are primarily based on circuit-switching. To date, these networks rely on Synchronous Optical NETWORK and Synchronous Digital Hierarchy (SONET/SDH) transport. Wavelength Division Multiplexing (WDM) is also deployed, but mainly for bundling SONET/SDH pipes into static point-to-point wavelength-based links. It is anticipated that carrier core networks will evolve to dynamic reconfigurable optical networks, the pillars of which are Wavelength Division Multiplexing (WDM), Optical Cross-Connect (OXC) nodes, and distributed control. Therefore, researchers are seeing the need for innovative techniques to groom multitudes of packet-based traffic streams into dynamic wavelength-based lightpaths. This need is emerging in both core and metro networks. Traffic grooming in optical networks, the topic of this book, has therefore become an important area of research over recent years [1–14].

In this chapter, we discuss how traffic grooming capabilities may be incorporated into OXCs. We focus on OXCs as primary optical network elements bearing in mind that grooming capabilities are usually embedded in Optical Add/Drop Multiplexers (OADMs) as well. The latter are important network elements especially in the metro domain. However, an OADM, be it static or reconfigurable (ROADM), can be regarded as a special case of an OXC for the purpose of our discussion of grooming and grooming switches. Unlike OXCs, an OADM has only one input optical-line port and one output port. Hence, it manages humble magnitudes of traffic compared to a small OXC, and therefore involves less grooming.

In the following section, we re-emphasize the case for traffic grooming in optical networks. This is followed by a summary of optical cross-connect types.

Then, we focus on grooming switches and shed some light on their architectures and their role in an OXC. This is followed by the classification of grooming abilities of grooming-capable OXCs.

2.2 Why Traffic Grooming?

Let us consider an optical core network carrying several independent traffic streams. We view each stream as a flow of IP packets entering the optical network at some ingress OXC node and exiting it from certain egress OXC node. Each stream represents a load which can be quantified in terms of a proportion of a wavelength bandwidth (λ). A stream may thus demand, say, 0.1λ , 0.5λ or 1.5λ worth of bandwidth to be transported. However, if the network is based on Optical Circuit Switching (OCS) [15] without any traffic grooming capability, each traffic stream will have to be assigned an integer number of wavelengths for transport. Hence, part of the link capacity will have to be unused.

Figure 2.1 depicts an example of four traffic streams emerging from an ingress node I and flowing in the same path through core nodes A , B , and C . The four streams exit the network at different egress nodes (E_1 through E_4) as shown in the figure. In an OCS network, an integer number of wavelength-based circuits is assigned to every stream. Let each traffic stream present a load that requires n wavelengths for a given buffer size and packet loss probability (PLP), where $n = 1, 2, 3, \dots$. Hence, a total of $4n$ wavelength channels would be required to transport all streams from node I to node C . Figure 2.2 shows a simplified schematic of this case at node I .

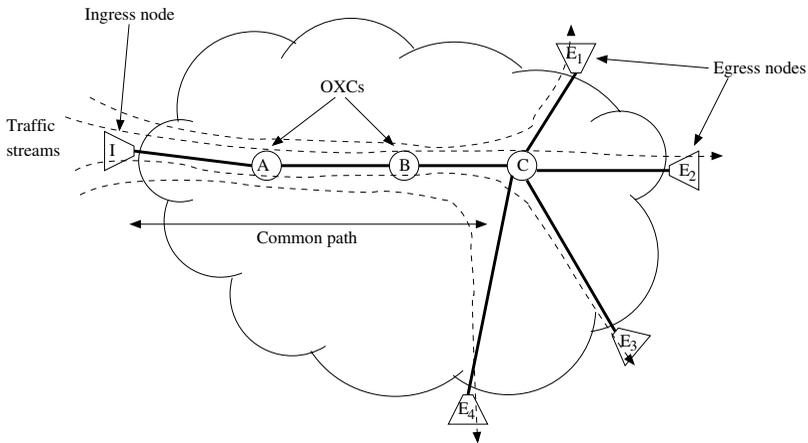


Fig. 2.1. Optical network carrying four traffic streams, Ingress node I , Egress nodes E_1 – E_4 , and core nodes A , B , and C are OXC nodes [10]

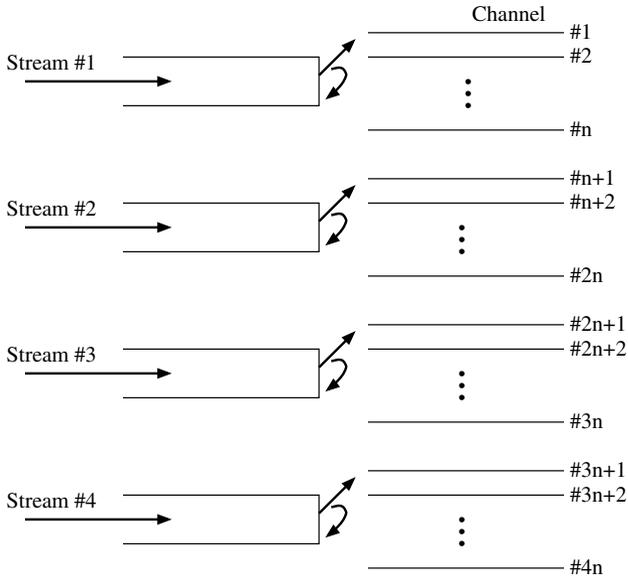


Fig. 2.2. Architecture of the ingress node I (Fig. 2.1) assuming a pure OCS-based network [10]

Now, if packet-processing capability is introduced at node I , traffic can be groomed in such a way as to achieve significant savings in bandwidth [10]. With grooming capability, the node can consolidate the same total magnitude of traffic into one set of channels as shown in Fig. 2.3. A total of m outgoing wavelengths would be required in this case, where m can be $\ll 4n$. In general, bandwidth savings depend on the number of traffic streams to be groomed, their statistical characteristics, buffer sizes, and packet loss probability.

A study was carried out to investigate the impact of traffic grooming on OXC-based networks assuming the above scenario. This study focuses on

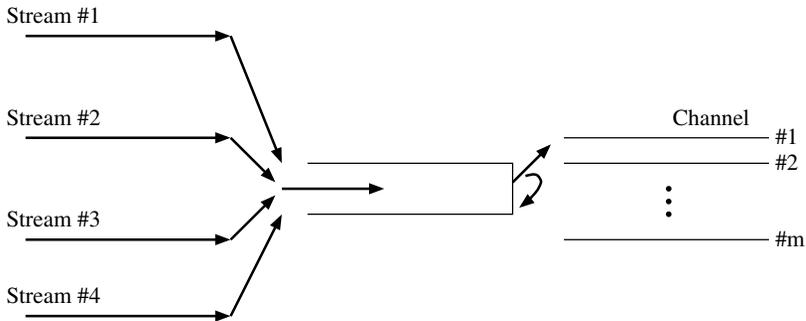


Fig. 2.3. Architecture of the ingress node I (Fig. 2.1) with packet grooming capability ($m < 4n$) [10]

quantifying potential bandwidth savings, which are enabled by traffic grooming, in terms of the statistical multiplexing gain (SMG). The results of this study show that SMG increases with increasing traffic variability and by increasing the number of traffic streams to be groomed. Also, SMG increases as packet loss probability (PLP) decreases [10].

Traffic grooming and statistical multiplexing enable several traffic streams to share a given wavelength channel(s). This lowers the number of wavelengths needed to transmit a given volume of traffic and improves bandwidth utilization and efficiency. By incorporating packet processing functionality in OXC nodes, traffic can be groomed in such a way that enhances bandwidth efficiency of the OXC. While the switch fabric is an important contributor to OXC cost, the number of ports (including Optical-to-Electrical-to-Optical -OEO- transponders, where applicable) is typically a more significant contributor to this cost. Also, the larger the number of wavelengths becomes, the larger the transmission plant cost becomes. Savings in OXC ports, OEO transponders, and in transmission system cost offset the cost of electronics needed for packet processing. An economic benefit can therefore be gained by traffic grooming. Of course, this is accomplished at the expense of extra equipment, processing, and control. Therefore, there is a trade-off between bandwidth utilization and OXC complexity.

2.3 Optical Cross-Connects

OXC's are advanced optical network elements that are capable of switching lightpaths, passing transit ones through, adding and dropping locally generated and terminated client-layer traffic, and configuring optical network topologies. When augmented with traffic grooming capability, OXC's can also groom client-layer traffic and carry out sophisticated bandwidth management tasks where complex network topologies and large numbers of wavelengths are involved. Therefore, OXC's are best suited for locations in the network where extensive bandwidth management is required, such as when several WDM transmission lines and many digital highways converge. OXC's are particularly useful for mesh topologies and to interconnect several WDM rings in core and metro networks. They enable the optical network to reconfigure in order to meet client layer needs and to get around node and link failures. They also help the network to be upgraded and maintained without service interruption [15, 16].

OXC's may be classified into four main types according to their switching capability [15]. The first type is the Fiber Cross-Connect (FXC), which switches all the wavelengths of an input fiber port to an output fiber port. This type is the simplest and least expensive of the four types. In effect, it acts like an automated fiber patch panel.

The second OXC type is the Wavelength-Band (wave-band) Cross-Connect (WBXC). In this cross-connect, traffic streams sharing the same end points

and having adjacent transmission wavelengths are grouped into bands which are switched all together. WBXC are more flexible than FXCs and provide more networking capability. This advantage comes at the expense of extra complexity and cost of the cross-connect.

The third type is the Wavelength Selective Cross-Connect (WSXC). It has the capability of switching individual wavelength channels simultaneously from any input port to any output port. It is a more able network element than the previous two types, but more complex and expensive too.

Finally, the Wavelength Interchanging Cross-Connect (WIXC) has the same switching capability of a WSXC, but adds to it the ability of wavelength conversion. As such, it is the most flexible, most complex and most expensive of the four types. Wavelength conversion is a useful feature to enhance the cross-connect blocking characteristics and to equip it with maximum networking flexibility. A lot of progress has occurred recently in wavelength conversion technologies and techniques. However, all optical wavelength converters have not reached yet the status of being widely available cost-effective commercial components.

Figure 2.4 provides a simplified schematic diagram of an OXC system that is equipped with traffic grooming capability. The cross-connect shown is of the WSXC type, but other types can be used to illustrate the same concept.

Optical Line Terminals (OLTs) are used in WDM point-to-point transmission systems [16]. At one end, channel wavelengths are multiplexed onto one fiber for outgoing transmission. At the other end, wavelengths are separated and converted to electrical domain where they are either delivered locally or passed over to another OLT. Typically, an OLT comprises transponders, a multiplexer/de-multiplexer (mux/demux) pair, and optical amplifiers. Transponders involve OEO-based wavelength conversion to adapt proprietary wavelengths to International Telecommunication Union (ITU)-standardized wavelengths. OLTs also extract the optical supervisory channel (OSC) which is used by the network for performance monitoring, Operation, Administration, and Maintenance (OAM), and control. These functionalities are carried out in the electronic domain. Several multiplexing technologies are commercially available today, the most important of which are gratings: bulk, arrayed waveguide gratings (AWGs), and fiber Bragg gratings (FBG); and filtering technologies, such as dielectric thin-film filters. After multiplexing, signals may undergo amplification. Erbium Doped Fiber Amplifiers (EDFAs) are typically used for this purpose. In principle, OLTs can be included as inherent part of an OXC. However, they are usually treated today as separate network elements.

The switch fabric is the core of an OXC. Optical switching devices are interconnected using various strategies to form a single-stage or multi-stage switch fabric. The interconnection can be based on Clos, Benes, banyan, tree, or any other architecture. Several fabric designs are possible with numerous switching characteristics. Switching devices may be based on one of several optical switching technologies [15]. Examples include

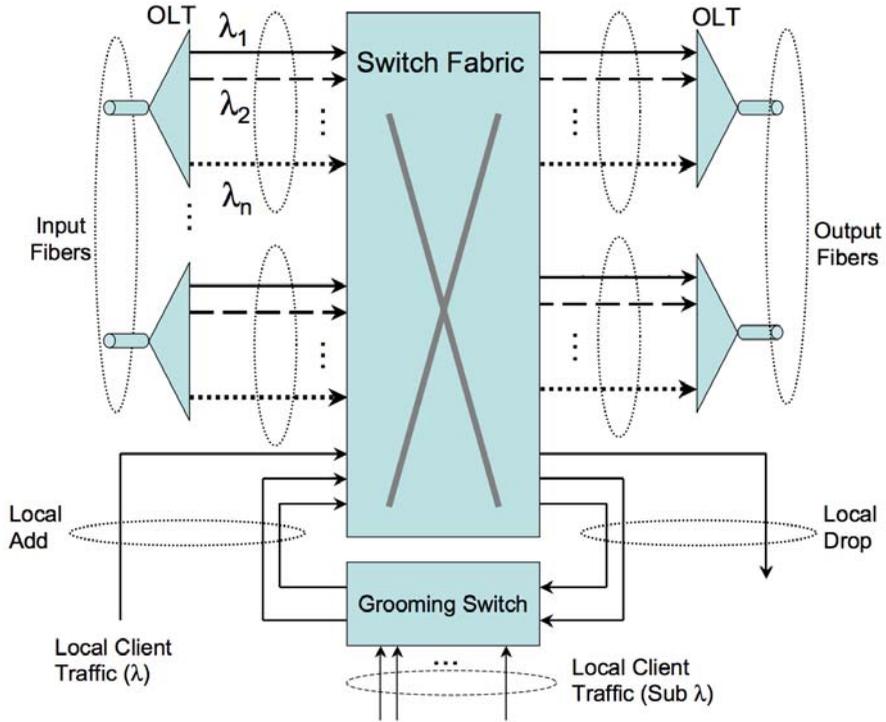


Fig. 2.4. Grooming-capable OXC

Electro-Optic (EO), Acousto-Optic (AO), Thermo-Optic (TO), Opto-Mechanical (OM), Liquid-Crystal (LC) based, Micro-Electro-Mechanical System (MEMS) based, Semiconductor-Optical-Amplifier (SOA) based, and Holographic optical switches [15]. Some technologies are more suitable than others for certain applications. It is possible to build the switch as one wavelength-independent fabric or to have a wavelength-partitioned architecture where a switch is provided for each wavelength [15].

While optical switching is a key to building true OXCs which can fully exploit the features of reconfigurable optical networking, some equipment vendors developed OXC systems that are based on electronic switching fabrics. In these systems, optical-to-electrical transponders convert input optical signals to the electrical domain and forward them to the electronic switch where they are directed to their designated output ports. The latter are equipped with electrical-to-optical converters whereby signals are converted back to the optical domain before transmission to next node(s). This type of cross-connects is referred to commercially as an OEO OXC or an opaque OXC (as opposed to transparent ones where signals remain in the optical domain throughout the switch fabric).

The grooming capability of the OXC of Fig. 2.4 resides in the module designated therein as the grooming switch. This module can indeed take a number of forms. It can be a SONET/SDH Digital Cross-Connect System (DCS), Next-Generation SONET/SDH (NG SONET/SDH) network element, IP router, Multi-Service Provisioning Platform (MSPP), or any other relevant telecommunication network element. Today, optical technologies have not matured enough to perform header processing functionalities in a commercial product. Therefore, this grooming unit is based on electronic technologies.

SONET/SDH is the standards and specifications that define the digital-transmission hierarchy of the public switched telephone network (PSTN). A SONET/SDH DCS is used to multiplex digital telephone circuits according to this hierarchy and is therefore considered as a tool to groom circuit-switching based traffic. Two main types of DCS exist, namely broadband DCS (B-DCS) and Wideband DCS (W-DCS). The main difference between the two types is that W-DCS goes deeper into the SONET/SDH digital hierarchy, down to the level of virtual tributaries. Packet-based traffic (ATM, IP, Ethernet, and others) is packed onto SONET/SDH circuits [17].

In recent years, a number of protocols were introduced to enhance the capability of SONET/SDH to carry packet-based traffic, and these developments led to NG SONET/SDH. Virtual Concatenation (VC), Link Capacity Adjustment Scheme (LCAS), and Generic Framing Procedure (GFP) enable SONET/SDH to carry and groom packet-based traffic more efficiently [17, 18].

MSPPs allow service providers to simplify their edge networks by consolidating a number of separate boxes into one intelligent access platform. They interface with various customer premises equipment, including TDM telephony, SONET/SDH, Ethernet, and broadband access. As such, they enable considerable traffic grooming capability.

In the following section, we discuss some details of typical grooming switches. For the purpose of this discussion, we assume the grooming switch to take the form of a generic packet switch/router and look into some architectures which are possible in this case.

2.4 Grooming Switches

The term *grooming switch* is used sometimes in the literature to refer to an OXC with traffic grooming capability (the reference here is to the entire cross-connect system). In this chapter, we restrict the definition of a grooming switch to the module within an OXC that is actually involved in traffic grooming. This terminology is more accurate and more consistent with industrial practice. The current state of the art, along with standardization and operational requirements, has led the telecom industry to separate between the OXC (be it transparent or opaque) and some auxiliary modules which are treated by many researchers as inherent part of the cross-connect itself. These modules include input/output ports, multiplexer and de-multiplexer

(mux/demux) units, OLT as a whole, and DCS. As such, carriers may mix and match equipment of various vendors and have flexibility in designing their networks subject to variable conditions.

As discussed before, several network elements can actually serve as a grooming switch. This includes traditional packet switches/routers. In this case, the switch can be designed based on one of several architectures [19–21]. Figure 2.5 depicts the simplest example where a general purpose workstation with multiple network interfaces, and suitable software, is used as a switch [19]. Here, a packet arriving from one interface can enter the memory where the CPU reads its header to determine on which interface the packet should be sent out. In practical shared-bus architectures, the general-purpose workstation is of course replaced by specialized equipment, and the bus is also different than traditional data buses. However, in the shared-bus architecture all ports/interfaces share a single common bus. In order to achieve non-blocking operation, the bus must operate at N times the port transmission rate (where N is the number of ports). Input ports take turns writing packets onto the bus at the full bus rate, typically in round-robin fashion. This architecture is inherently limited in terms of bandwidth and scalability.

Another example is the ring switch architecture where the bus is replaced by a ring. Each port has an interface to the ring. All input ports share the ring using a contention scheme (such as the token passing scheme). The architecture resembles the shared bus in many aspects, but the ring architecture supports transmission concurrency and can therefore have higher port bandwidth.

In the crossbar switch architecture, a number of *crosspoints* form the switching matrix. When the crosspoint in position (x, y) is closed, data signals

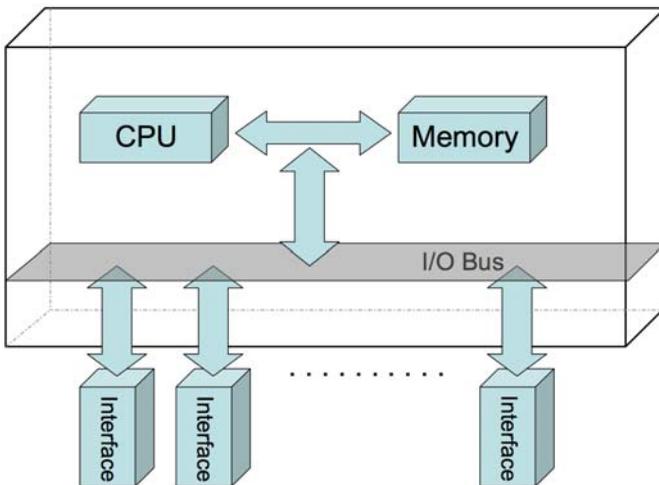


Fig. 2.5. A simplified shared-bus packet switch

can be transferred from input x to output y . Data transport typically take place in cycles and a control system reconfigures the switch fabric each cycle in order to coordinate packet flows among several input and output ports.

Some switch fabrics are known as *self-routing* fabrics. In these fabrics, a special self-routing header is appended to the packet by an input port after it determines which output port the packet should go to. This extra header is removed before the packet leaves the switch. Self-routing fabrics are often built using a large number of simple 2×2 switching devices interconnected in standard switch fabric configurations, such as the Banyan network [19].

In the shared-memory architecture, a pool of buffers is shared among input and output ports (instead of port dedicated buffers which can be used with previous architectures). A mechanism is needed of course to transfer data from input ports into shared memory and from the latter to output ports. Two crossbars are typically used: one between inputs and the memory banks and the other between the memory banks and outputs.

Single-stage switching fabrics are limited in scalability. Therefore, scalable switches are built using multi-stage architectures. This is a trend which has been rising in the routing industry over recent years. Multi-stage architectures differ depending on how they interconnect single-stage building blocks.

Finally, switching-fabric integrated circuits are built using different technologies, i.e., Field-Programmable Gate Arrays (FPGAs) and Application-Specific Integrated Circuits (ASICs) provide two main design approaches. The designer's choice of any of these two approaches depends on cost, performance, scalability, blocking characteristics, configurability, signal integrity, and many other considerations.

FPGAs offer the flexibility of reprogramming the switch as needed and enable designers to augment the switch with additional functionalities. ASICs offer cost-effective solution for high-volume designs which are cost sensitive and require specific functions.

The telecommunication industry is now also familiar with a third option, which is Application-Specific Standard Products (ASSPs). These are available as off-the-shelf components. Therefore they facilitate faster time-to-market than ASICs. It is argued also that they provide better switching performance than FPGAs. Nevertheless, they do not have the same ability to integrate additional functionalities like FPGAs and ASICs.

2.5 Classification of Grooming-Capable OXCs

We have seen how the traffic grooming functionality may be incorporated in an OXC. In [8], grooming capable OXCs are classified according to their overall architecture and to how they are placed in the network into four main types: single-hop-grooming OXCs, multi-hop partial-grooming OXCs, multi-hop full-grooming OXCs, and light-tree based source-node grooming OXCs. In the following, we briefly explain this method of classification.

In a single-hop grooming OXC, low-data-rate client traffic can be multiplexed onto wavelengths and all traffic that is carried over a given wavelength channel is switched to the same destination port. No switching capability, per se, is provided. The grooming unit in this case is a traffic aggregation unit. As such, the single-hop grooming scheme has limited grooming capability in that it can only groom traffic from the same source node to the same destination node.

A multi-hop partial-grooming OXC, on the other hand, has two switch fabrics (as the case in Fig. 2.4). The main switch can be an OCS optical switch or an OEO switch for wavelength-based circuits. The second switch is the electronic grooming switch for low-speed client traffic. This OXC has more grooming power and flexibility than the single-hop type and client traffic can be transferred from one wavelength to another. Only few wavelength channels need to be delivered to the grooming fabric. The number of ports connecting both switching fabrics, however, determines the grooming capability of the cross-connect.

In a multi-hop full-grooming OXC, every input wavelength is demultiplexed to its constituent electronic traffic components. One large electronic switch fabric is used, in this case, in lieu of the two fabrics in Fig. 2.4. The switch can mix and match input traffic components as desired and can consolidate them as efficiently as possible. Then, all traffic is wavelength multiplexed onto outgoing wavelengths. For example, any SONET circuit can be switched from a given input time slot(s) of a wavelength channel to another output slot(s) of a different wavelength. Traditional digital switching technologies (combination of space and time switches) can be used to implement the switch fabric in this case. This fully electronic approach, however, involves a great deal of switch complexity and is not scalable beyond certain point.

All in all, the multi-hop grooming schemes may groom traffic from different source nodes to different destination nodes.

Finally, a *light tree* is designed to support multi-cast applications in optical network [8, 22]. In a light tree, an OXC which originates the traffic is referred to as a root node whereas OXCs which terminate traffic are considered leaf nodes. For multicasting, the cross-connect has to be capable of generating multiple copies of an input signal and direct them towards outputs. In the optical domain, this can be accomplished by passive splitters. In the electrical domain, the signal can be copied into memories. By setting up a light-tree, a root node can groom several traffic streams together onto one wavelength channel. The light-tree-based source node grooming scheme can groom traffic from the same source node to different destination nodes.

While the multi-hop full-grooming approach can provide best performance in terms of resource utilization and blocking characteristics, it can have scalability limitations. The multi-hop partial grooming approach offers reasonable alternative when full grooming is not necessary in each and every node. In

any practical network setting, however, one should assume different nodes with different grooming capabilities.

Finally, intelligent algorithms are necessary for grooming-capable OXCs to manage their traffic as efficiently as possible. This topic is discussed in more detail elsewhere in this book.

References

1. S. Thiagarajan and A. K. Somani, "Capacity fairness of WDM networks with grooming capability," *Opt. Networks Mag.*, vol. 2, no. 3, pp. 24–31, May/June 2001.
2. E. Modiano and P. J. Lin, "Traffic grooming in WDM networks," *IEEE Commun. Mag.*, pp. 124–1299, July 2001.
3. R. Ranganathan, J. Berthold, L. Blair, and J. Berthold, "Architectural implications of core grooming in a 46-node USA optical network," in *Proc. IEEE/OSA OFC '02*, Mar. 2002, pp. 498–499.
4. R. Dutta and G. N. Rouskas, "Traffic grooming in WDM networks: past and future," *IEEE Network*, pp. 46–56, November/December 2002.
5. K. Zhu and B. Mukherjee, "A review of traffic grooming in WDM optical networks: architectures and challenges," *Opt. Networks Mag.*, vol. 4, no. 2, pp. 55–64, March/April 2003.
6. L. B. Sofman, A. Agrawal, and T. S. El-Bawab, "Traffic grooming and bandwidth efficiency in packet and burst switched networks," in *Proc. of the Applied Telecommunications Symposium (ATS 2003)*, part of the Advanced Simulation Technologies Conference (ASTC 2003), pp. 3–8, Orlando, Florida, March 30–April 3, 2003.
7. H. Zhu, H. Zang, K. Zhu, and B. Mukherjee, "A novel generic graph model for traffic grooming in heterogeneous WDM mesh networks," *IEEE/ACM Trans. on Networking*, vol. 11, no. 2, pp. 285–99, April 2003.
8. K. Zhu, H. Zang, and B. Mukherjee, "A comprehensive study on next-generation optical grooming switches," *IEEE J. Select. Areas Commun.*, vol. 21, no. 7, pp. 1173–86, September 2003.
9. L. B. Sofman, T. S. El-Bawab, and A. Agrawal, "Effect of traffic variability on statistical multiplexing gain in a continuous bufferless traffic model," in *Proc. SPIE*, vol. 5244, pp. 122–30, Performance and Control of Next-Generation Communications Networks, part of SPIE International Symposium ITCOM 2003 (Information Technologies and Communications), September 7–11, 2003, Orlando, Florida, USA.
10. A. Agrawal, L. B. Sofman, and T. S. El-Bawab, "On the enhancement of bandwidth efficiency by traffic grooming in optical cross-connect based networks," in *Proc. SPIE*, vol. 5247, pp. 196–202, Optical Transmission Systems and Equipment for WDM Networking II, part of SPIE International Symposium ITCOM 2003 (Information Technologies and Communications), September 7–11, 2003, Orlando, Florida, USA.
11. S. Huang, R. Dutta, and G. N. Rouskas, "Traffic grooming in path, star, and tree networks: complexity, bounds, and algorithms," *IEEE J. Select. Areas in Commun.*, vol. 24, no. 4, April 2006.

12. C. Xin, B. Wang, X. Cao, and J. Li, "Logical topology design for dynamic traffic grooming in WDM optical networks," *IEEE J. Lightwave Technol.*, vol. 24, no. 6, pp. 2267–75, June 2006.
13. A. E. Kamal, "Algorithms for multicast traffic grooming in WDM mesh networks," *IEEE Commun. Mag.*, pp. 96–105, November 2006.
14. S. Huang and R. Dutta, "Dynamic traffic grooming: The changing role of traffic grooming," *IEEE Commun. Surveys & Tutorials*, vol. 8, no. 4, 4th Quarter 2006.
15. T. S. El-Bawab, "Optical Switching," Springer, July 2006, ISBN: 0-387-26141-9.
16. R. Ramaswami and K. N. Sivarajan, "Optical Networks: A Practical Perspective," 2nd Ed., Morgan Kaufmann, Academic Press, 2002.
17. W. Goralski, "SONET/SDH," 3rd Ed., McGraw-Hill, Osborne, 2002.
18. P. Bonenfant and A. Rodriguez-Moral, "Generic framing procedure (GFP): the catalyst for efficient data over transport," *IEEE Commun. Mag.*, vol. 40, no. 5, pp. 72–9, May 2002.
19. L. L. Peterson and B. S. Davie, "Computer Networks: A System Approach," 4th Ed., Morgan Kaufmann Publishers, 2007.
20. J. Bergen, "Selection of proper switch fabric tied to ultimate application," *Lightwave*, pp. 130–131, July 2000
21. G. Parulkar, "Harnessing the power of the optical internet," *Lightwave*, pp. 132–133, July 2000.
22. L. H. Sahasrabudde and B. Mukherjee, "Light-trees: optical multicasting for improved performance in wavelength-routed networks," *IEEE Commun. Mag.*, vol. 37, pp. 67–73, February 1999.

Control Plane Support

Slobodanka Tomic

3.1 Introduction

Novel transport networks with opto-electronic multiplexing and switching can use static or dynamic traffic grooming methods to cost-efficiently resolve the mismatch between the high capacity of wavelengths and the low bandwidth requirements of predominantly IP or Ethernet services. The multiplexing and switching at different granularity layers in optical or digital hierarchy can be either *selectively deployed* to achieve cost-efficient network design for anticipated multi-granular traffic (static grooming) or *selectively used* to maximize network throughput for dynamically changing traffic (dynamic grooming) [1] [2]. A distinguishing feature of grooming networks is *flexibility* – only at the lowest layer in the hierarchy a network has a static topology; at each higher layer, the topology is virtual and can be changed, i.e., *engineered*, by setting up or releasing end-to-end connections in the underlying layer. To exploit this flexibility when the traffic is dynamically changing, topology engineering requires an automatic process that combines traffic monitoring and fast network re-configuration. Today, an answer to this requirement provide the Generalized Multiprotocol Label Switching (GMPLS) [3] framework of the Internet Engineering Task Force (IETF), and Automatic Switched Optical Network (ASON) framework of the International Telecommunication Union-Telecommunication Sector (ITU-T) [4]. These two standardization frameworks define a new paradigm for service provision and resource management in multi-layer networks, the core of which make:

- *A transition from a traditionally centralized management-plane (MP) approach to network operation, towards a control-plane (CP) approach.* CP-approach is based on automatic and distributed control functions, and therefore it promises higher scalability and robustness, and lower operational costs. In fact, OPEX reduction in the order of 50% can be expected for most telecom operator models [5].

- *A unified resource control model applicable to multiple layers.* In this respect, GMPLS proposes a model with a five-layer hierarchy of switching capabilities for *IP-over-WDM* converged infrastructure, including packet switching (PSC), layer-2 switching (L2SC), time division multiplex switching (TDM), lambda switching (LSC), and fiber switching (FSC) [6]. ASON defines a model for optical transport network technologies with a “rich” TDM switching layer (*IP-over-OTN*), including synchronous digital hierarchy (SDH) [7], synchronous optical network (SONET) [8], and optical transport network (OTN) [9]. The different scope and therefore the mismatch between the ASON and GMPLS models becomes obvious when the switching hierarchies are compared. The optical channel (OCh) switching capability of the OTN hierarchy is mapped into the lambda switching capability (LSC) of GMPLS; on the other hand, switching at different path layers of SDH/SONET, as well as at the digital path layers of the OTN hierarchy map into just one Internet switching capability type being TDM. This has implications on the model, for example, a topological representation of a TDM layer in GMPLS, a TDM Region [6], includes different SDH/SONET/OTN interfaces.

The two frameworks identified the core CP components as classified within three major processes, which should support dynamic and automatic provision, and eliminate the “human factor” often found responsible for misconfigurations and provision delays. These processes are:

- Connection and call control (signaling).
- Distribution of network information and path calculation (routing).
- Automatic resource discovery and inventory.

Within this scope, the ASON framework particularly focuses on control plane architecture and protocol neutral functionality, including special purpose controllers and their relationship. GMPLS, on the other hand, defines protocols for realization of CP-functions. GMPLS extends two MPLS routing protocols, namely the Open Shortest Path First extended for Traffic Engineering (OSPF-TE) [10], and the Inter-System to Inter-System extended for Traffic Engineering (ISIS-TE) [11], and two signaling protocols, the Reservation Protocol extended for Traffic Engineering (RSVP-TE) [12] and the Constrained Label Distribution protocol (CR-LDP) [13]. Signaling extensions [14] support distributed configuration of circuits of different types, i.e., generalized label switched paths (LSP). This includes reservation of link resources allocated to LSPs (e.g., wavelengths or a time slots), realized by means of the link-local distribution of generalized labels, and configuration of switching fabrics, along the path. Routing extensions [15] support distribution of TE attributes for GMPLS TE links, e.g., by periodic advertisements, as a support for distributed path calculation. Link Management Protocol (LMP) [16] supports the discovery and configuration of the transport and control plane interfaces and their mapping. Within GMPLS, the current work focuses on specific requirements of IP-over-OTN grooming networks, which are potentially superior to

Table 3.1. Grooming requirements supported within GMPLS framework

Grooming Requirements	GMPLS / PCE Support
	Routing / Path control
<ul style="list-style-type: none"> • A model for multiplexing and switching capability for multilayer TE links • A model for internal multiplexing capability of grooming nodes • Multi-layer routing • Virtual topology management 	<ul style="list-style-type: none"> • Switching capability descriptor (ISCD) of an advertised external interface pertaining to a TE link [10, 11]. • Adaptation capability descriptor (IACD) of an advertised internal interface [26]. • Hierarchical information aggregation / advertisement [30]. • Hierarchical path calculation with PCE [32]. • Model of a virtual link [26].
	Signaling / Call and Connection Control
<ul style="list-style-type: none"> • Definition of a complex signal structure and identification of each signal component within the connection / call • Call control 	<ul style="list-style-type: none"> • LSP signal attributes, generalized labels [34, 35]. • Support for VCAT/LCAS discovery [36]. • Call identification, distributed setup with crank-back [37].

IP-over-WDM networks in terms of scalability, flexibility, and robustness as shown in [17]. The objective is to fill in the gaps between the generic functionality supported so far, and the required functionality, often assumed in theoretic traffic grooming studies. This is an essential task, because the applicability of the traffic grooming methods, in particular, multi-layer routing algorithms using specific network graph models, such as those proposed in [1, 18], largely depend on the availability of the network state information used to construct the routing graph. The required availability can be assumed as fully given with a centralized management system; due to scalability issues, however, the distributed CP-based provision must rely on aggregated information, and the appropriate aggregated representation of resources is needed. Table 3.1 summarizes the grooming features addressed within the GMPLS framework that are briefly described in this chapter.

This chapter reviews the GMPLS/ASON support for CP-based service provision in grooming networks with IP-over-OTN infrastructure. Section 3.2 briefly introduces optical transport network (OTN) as an advanced grooming platform based on opaque (translucent) optical technology. Section 3.3 introduces the framework within which the control plane mechanisms supporting grooming are considered. The components of the framework are addressed in Sections 3.4–3.8, which focus on CP-support for resource modeling, discovery, topology advertisement, resource allocation, call control, virtual network control, and services. Section 3.9 concludes the chapter.

3.2 Grooming in Optical Transport Network (OTN)

Today, opaque optical networks, which implement the optical channel by means of a digital framed signal with digital overhead, outperform transparent networks, because they are able to guarantee accurate assessment of the quality for the *digital* client signals [19].

The optical transport network (OTN) [9] enhanced with the ASON/GMPLS control plane can be considered as an important traffic grooming platform. OTN uses efficient methods to monitor and guarantee service quality based on Forward Error Correction (FEC) and the embedded control overhead; it exploits flexible client signal mapping with the Generic Framing Procedure (GFP), i.e., the digital wrapper, and it supports application-adaptable bandwidth management. In OTN, the bandwidth of a wavelength at the digital layer is controlled by flexible switching, multiplexing, and mapping of 2.5G (Optical Data Unit – ODU1), 10G (ODU2), and 40G (ODU3) signals. The standard supports *variable adaptations*: an interface can be dynamically configured with one specific type of mapping, e.g., either 4 ODU2 or 16 ODU1 be mapped onto one ODU3. Based on GFP and the support for SDH clients, OTN can accommodate services of practically any granularity. Bandwidth efficient accommodation of different client signals is enabled with Virtual Concatenation (VCAT) [20] and Link Capacity Adjustment Scheme (LCAS) Protocol [21]. VCAT is a layer 1 inverse multiplexing technique for OTN, SONET, and plesiochronous digital hierarchy (PDH) component signals. By means of VCAT multiple paths in the server layer can be combined into an aggregate link in the client layer with virtually adjustable bandwidth. VCAT/LCAS provide a number of new network features such as flexible concatenation of containers at the right granularity (e.g., 100 Mb/s Ethernet is equal to VC-3-2v, or STS-1-2v). Component links may be possibly realized over multi-hop paths taking different routes. VCAT can support bandwidth on demand and IP traffic engineering by enhancing the capacity of the link without changing the topology in the IP layer. A new component link can be routed and set up over new available path before the old component is released, which is characterized as painless re-grooming [22]. Complemented with LCAS, VCAT can support realization of new forms of protection/restoration and graceful degradation based on dynamic repair. LCAS can dynamically change the capacity of an established VCAT group, automatically decreasing the capacity in case of a group member failure, and increasing the capacity after the repair. Summarizing the major grooming enablers in the transport plane, which may be exploited by the control plane, are flexible mapping of services, variable adaptations which may be dynamically activated, dynamic modification of link/call bandwidth achieved by adding or releasing component signals, and virtual concatenation of component signals with path-independent routing.

3.3 Control Plane Support for Traffic Grooming

The control plane support for traffic grooming includes mechanisms and protocols for resource-optimal configuration of flexible grooming capability and for optimized traffic routing. Figure 3.1 maps the required CP functionality into the framework of three functional blocks related to network resource control,

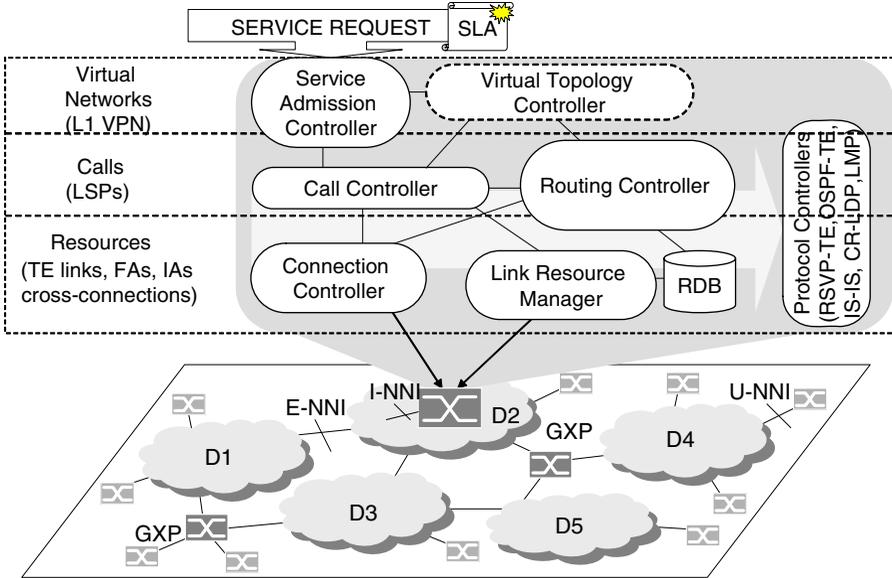


Fig. 3.1. CP-support framework

call control, and virtual network control. So far, ASON control plane concepts and GMPLS protocols focus mainly on functionality needed for resource control, and call routing and control.

Figure 3.1 also shows a general network scope for CP-based operation which is the multi-layer multi-domain network, operated by different operators, and providing services of different granularity. The reference ASON architecture currently features three types of characteristic CP-interfaces, namely user-to-network interface (UNI) between a service user and a provider domain, internal network-to-network interface (I-NNI) between the network elements within one carrier domain, and external network-to-network interface (E-NNI) between the network elements of different carrier domains. These interfaces differ in the level of the network information exchange they may support. While internal I-NNI supports full information sharing, UNI and E-NNI often require information hiding. Conceptually these interfaces relate to GMPLS models of the control plane integration which are the *overlay model* with no routing information exchange, the *peer model* with full sharing (and integration) of control plane instances and the *augmented model* with the aggregated information exchange [3]. In Fig. 3.1, we show the standard multi-domain architecture enhanced with challenging new elements, referred to as GMPLS-based exchange points (GXP), [23], which enable both flexible inter-domain connections between different administrative domains and customers and their GMPLS-based control. Sections that follow review functionality inherent to different blocks of the framework.

3.4 Resources

Due to scalability and trust issues, distributed CP-based provision must rely on aggregated information, and the appropriate aggregated representation of resources is essential. GMPLS and ASON developed specific network resource models by using two different set of architecture tools. The ASON model is based on the generic functional model for transport networks (G.805), [24], and is inherently “topological”. A network is modeled with a number of layer topologies in a client–server relationship. The example in Fig. 3.2 illustrates the ASON model for network with an optical channel switching layer and the OC-3 switching layer. The relationship between GMPLS and ASON models shown in the table is briefly covered in the discussion that follows.

Each layer topology provides switched service between client access points over interconnected sub-networks (SN), where the switching matrix is the smallest SN. The interconnection between two SNs is a link with one or multiple link connections. The end-points of link connections are sub-network connection points (SNP). Logically coupled SNPs form a pool (SNPP), e.g., all OC-3 within a OC-192 multiplex are coupled within one pool, and consequently a transport capacity between two SNPPs of two sub-networks, or between an SNPP on a sub-network and an access SNPP (the latter being the case in Table 3.2) is referred to as an SNPP link. ASON service is an SNP trail that connects two client access points, and is routed over a number of SNP link connections or sub-network connections. The association between an SNP in a server layer (optical channel - λ) and an SNPP (of OC-3 SNPs mapped within OC-192) in the client layer represents a specific termination and adaptation between the two layers in the multi-service switching node.

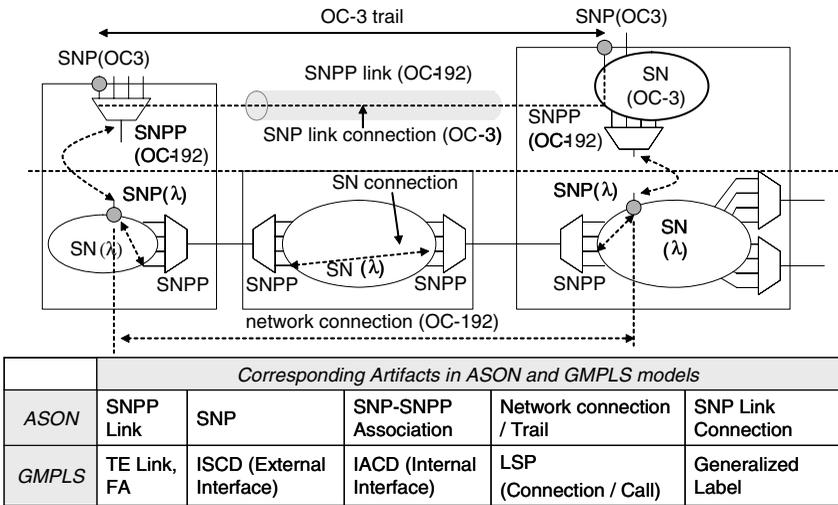


Fig. 3.2. Example of ASON modeling and ASON and GMPLS relationship

Table 3.2. Multiple interface switching capability descriptors.

Photonic XC	Hybrid XC	ISC Descriptor	ISC Type	Encoding	Max BW	Min BW
X	X	ISCD-1	LSC	SDH	STM-64	–
–	X	ISCD-2	TDM	SDH	STM-64	VC-3
–	X	ISCD-3	TDM	SDH	STM-64	VC-4

With variable adaptations that can be flexibly activated, several different associations can exist in the same time. ASON is particularly designed for discrete bandwidth networks: an *atomic transport resource* is an SNP link connection. Obviously, in terms of routing requirements, advertising SNP link connections would not scale, therefore SNPP link would be advertised as an aggregate characterized by the number of available SNP LCs. An association SN-SNPP uniquely identifies an *atomic grooming resource* which must be allocated along the multi-layer routed path, and therefore, should also be advertised to support distributed path calculation.

As already mentioned, the ASON resource model is based on SDH/SONET/OTN functional modeling and extensions to GMPLS were needed to account for a rich TDM layer. In this context, several ASON subnetwork points (SNP) and/or point pools (SNPP) map to an interface, a termination of a GMPLS TE link. A GMPLS interface can either be a simple interface (SNP) or a combination of component interfaces (SNPP). To cope with the multiple parallel interfaces between two nodes, the interfaces can be unnumbered [25]. For some technologies, a component interface is an atomic resource, e.g., a wavelength, and is addressed with a generalized label (corresponding to an SNP LC). For SDH/SONET and OTN digital hierarchy, a component interface has a multiplex of smaller discrete resources, each of which is described with essentially more complex generalized label. One interface can therefore represent several SNPs or SNPPs at different ASON layers, each of which can be represented with an Interface Switching Capability Descriptors (ISCD). ISCD includes a switching type (LSC or TDM) and maximum and minimum reservable bandwidth available, the latter depending on the adaptation configured at the interface. Examples of ISCD usage are given in Table 3.2. An interface of a photonic cross-connect is fully described with LSC switching (ISCD-1). In addition to ISCD-1, an interface of a hybrid (SDH-WDM) cross-connect is described also with ISCD-2 and ISCD-3 (TDM switching) defining VC-4 and VC-3 as a minimum reservable bandwidth which can be allocated at this interface. However, if VC-4 is allocated, ISCD-2 will disappear from the ISCD list.

The GMPLS interface model becomes rather complex when used for representing the capabilities of multi-service switching nodes. The concept of a GMPLS TE link is not meant to be used to model internal multiplexing capability of nodes, although in fact, SN-SNPP association could be modeled as TE link with one interface in a client layer and the other in a server layer. Without

such modeling, the available grooming capacity of a node cannot be used to calculate a TE path for a service. Within GMPLS, a concept of the Interface Adaptation Capability Descriptor (IACD), which is a new attribute defined for an internal interface which could be advertised with GMPLS routing protocols, is currently being evaluated [26], and also experimentally verified [27]. IACD describes an interface in terms of two ISC Descriptors (for both ends of the interface or link) with the standard bandwidth encoding depending on the ISC type. Both ASON and GMPLS define a client server relationship between two layers in hierarchy. In the ASON model, the link in the client layer is realized over a network connection in the server layer. A client layer link is also a part of an end-to-end network connection supporting the client layer trail. A GMPLS client-server relationship associates a dynamic TE link in a client layer, referred to as a Forwarding Adjacency (FA), and a label switched path (LSP) in a server layer, established between two interfaces that implement client and server switching types. By using the newly established FA, a client layer LSP can embed itself within the server layer LSP. For routing scalability, all parallel resources with the same TE capability could be bundled into, and advertised as, one bundled TE-link or FA [28]. Each TE link or FA is also assigned to some shared risk link group (SRLG): all links that fail together are assigned the same SRLG value. Also each link is assigned a protection level such as extra traffic, unprotected, shared, and dedicated 1:1 or 1+1.

3.5 Resource Control

The functionality of the control plane processes within the functional block RESOURCES can be shortly summarized with “*discover, advertise, reserve*”, the issues of which are described in the following sub-sections.

3.5.1 Resource Discovery

A discovery process is needed for automatic correlation of interfaces between either two neighboring nodes or neighboring technology layers within one node. Within the discovery process both transport connections and control channels are discovered and verified. A central component in the ASON discovery process is the Link resource manager (LRM), which maintains the local inventory of links and updates a local routing database (RDB) with configured TE links. The routing database can be configured also through a management system. The complexity of the procedures needed to create a TE, best illustrated in the MIB document [29], however, offers strong motivation for automation of the discovery process. The discovery process establishes the binding between the management plane (MP) names of the transport resources (modeled with G.805) and the control plane (CP) names of the same resources (SNPs) at one node. It also logically binds the MP names and the corresponding CP names of potential link connections (SNP-SNP), between the neighboring nodes. A potential SNP-SNP link connection

becomes an actual LC when a corresponding flexible adaptation is activated. In the GMPLS framework, the link management protocol (LMP) [16] supports the resource discovery process with four basic functions for a node pair. The control channel management establishes and maintains connectivity between adjacent nodes. The link verification procedure verifies the physical connectivity. The link summary messages are exchanged to correlate link properties between adjacent nodes, first when the node is being brought up and then periodically when a link is up and not in the verification procedure. Finally, LMP provides a mechanism to isolate link and channel failures in both opaque and transparent networks, independent of the data format.

3.5.2 Resource Advertisement

The advertisement of network topology maintained within the routing data base (RDB) is a process which enables distributed path calculation for requested services. For the purpose of scalability and data hiding, ASON proposes that the advertised TE links (and internal interfaces), be organized within multiple routing areas (views) which filter data based on different operational constraints. The filtering could be based on a technology layer (e.g., TDM, LSC), administrative aggregation level (inter-domain, intra-domain), shared risk link group (SRLG) identifier, or some other type of administrative grouping. Filtering can also be subject to operational constraints related to a type of an interface (e.g., UNI, I-NNI, E-NNI) and the CP-integration models (overlay, peer, augmented). Regarding the organization of the CP-functions, the advertisement process at a particular grooming node is responsible for a number of routing controllers logically associated with different routing areas and collocated at the same node. For example, Fig. 3.3 shows the hierarchical

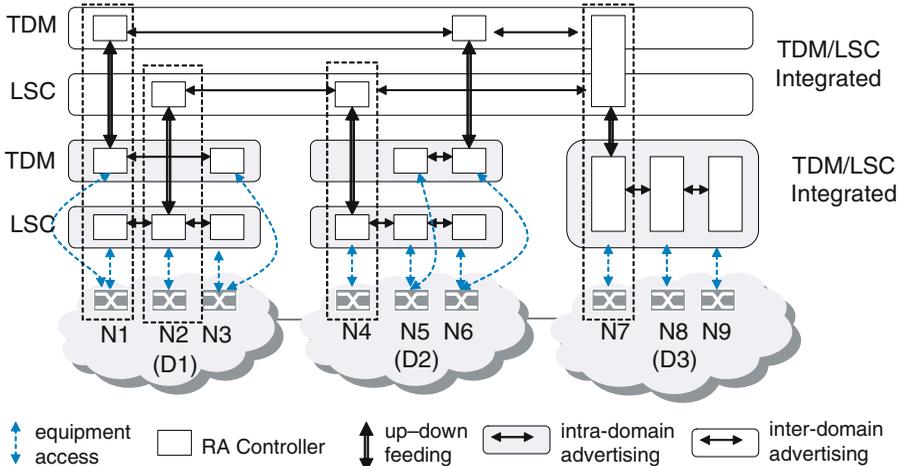


Fig. 3.3. Hierarchical routing areas in multi-domain TDM-LSC network

architecture of the advertisement process in a network of three interconnected domains that implement LSC and TDM switching.

In domains D1 and D2, advertisement of TE attributes of LSC and TDM layer is separated (overlay model) and in D3 integrated (peer model). The controllers at nodes (N1, N6 and N7) and (N2, N4 and N7) take part in the inter-domain information exchange at the TDM and LSC layers, respectively. These controllers exchange routing information by feeding-up or feeding-down data, and communicate with their (intra-domain) peers at other nodes. The vertical information flow may support different routing paradigms, e.g., for hierarchical routing only feeding-up is necessary because the path is calculated top-down, starting from the highest possible level. For source-based routing, feeding down is needed as the routing decisions are made at the lowest level of the routing hierarchy. Driven by ASON routing requirements, proposals extending the OSPF-TE and ISIS-TE protocols for hierarchical operation are currently under study [30], with feed-up and feed-down capabilities between the hierarchically organized OSPF-TE and ISIS-TE areas. An important aspect related to TE link advertisements in GMPLS networks is that the advertised changes not only reflect the update of TE link attributes due to resource allocation (reduced residual bandwidth), but they also show the creation or deletion of dynamic TE links (FAs). Currently, a FA is always allocated with a bandwidth. However, with *virtual links*, which can have no bandwidth at all, this dynamic could be decreased. The extensions for virtual links are also under study [26]. Link bundling and advertisement of bundled links is also a challenging task because of the parallel links with different TE features. Therefore, taking into account the dynamics and the nature of advertised changes, the process of advertising in grooming networks needs new consideration. The latency and accuracy of data updates, the issue of information hiding, the impact of distributed resource reservation and others, are related issues of interest, as also reflected in recently published studies, e.g., [31]. Within the IETF, a Path Computation Element Framework defines a new architecture [32] for a routing process decoupled from GMPLS protocols. Similar distributed trusted routing service is proposed in [33] to either provide a path upon request or provide for data filtering functions between the distinct control planes.

3.5.3 Distributed Resource Reservation

Based on the advertised information, a service path can be calculated; however, particular resources (generalized labels) on TE links along the path can be selected and allocated only in a link-local hop-by-hop resource reservation process. In GMPLS, resource reservation is performed within a distributed signaling session along the path of an LSP. In an RSVP-TE session, path message is sent from the the LSP ingress controller to the LSP egress controller, carrying objects which define end-to-end connection requirements, such as the signal type and required bandwidth, and objects that have link-local meaning

and support resource/label selection process. The explicit route object (ERO), which is provided by the path calculation controller at the source of an LSP, is, in general, also a link-local attribute and could be changed at each intermediate controller that is also involved in path calculation. On each link, the label selection is supported by *the explicit label*, where an upstream node provides a label preference to a downstream node, to be accepted or rejected. The upstream node can start configuring its hardware with the proposed label in advance, which can, in case of accepted label, reduce the setup latency, e.g., when restoration LSPs need to be rapidly established. With a *label set* object, an upstream node can suggest several preferred/acceptable resources for connection. In the optical domain, by means of a label set, a common available wavelength along the whole path may be determined. An *Upstream label* is used in the allocation of bi-directional links for bi-directional LSPs, with both directions following the same path and having the same traffic engineering attributes, including the protection and restoration level. The two directions of a link are allocated at the same time and, therefore, the setup latency and the control overhead are equal to those of a unidirectional LSP setup.

GMPLS supports allocation of time slots within SDH/SONET and OTN hierarchies encoded in generalized labels which describe both the multiplexing structure and the position of the signal in the multiplex [34, 35]. With these extensions, one SDH/SONET/OTN LSP can require allocation of several labels, by specifying a number of multiplex signals and the structure of this elementary multiplex given with a number of elementary signals either contiguously or virtually concatenated. With fixed adaptations, a signaled request may result into a number of LSPs set up first on the lower layers (e.g. ODU3) and then on the requested layer (ODU1). However, configuration of flexible adaptations is not supported in a straightforward manner as a label request does not specify a multiplex structure. Even assuming that the internal adaptation capabilities are advertised and used to create a routing graph for ODU1, ODU2 and ODU3 layer, transforming a request into a specific ERO and activating flexible adaptations is a challenging task which is still not fully supported.

3.6 CALL Control

In ASON, a call is a logical association between two customers and is hence the simplest service. A connection, on the other hand, is the realization of a call. A call can be associated with zero or a number of connections that may be established, released, modified, and used for different purposes (e.g., normal operation or protection). In particular, in transport networks where restorability and bandwidth adaptability are service *features*, the control plane support for bandwidth modification, and protection/restoration of a call are of major interest. This is identified as a major feature and extensions to GMPLS which support a call with no connections are under consideration. Regarding

bandwidth modification, the aggregation of multiple connections within one call needs to be supported within both the transport and control plane. Figure 3.4 illustrates this issue. A call can be provisioned over contiguously concatenated signals over a path (N1, N2, N3, N4) or as a virtually concatenated signal split over different paths between N5 and N6. If the availability of source and sink VCAT functions can be discovered, the path-independent routing of call components could be supported. In grooming networks, work on enhancing GMPLS/G.ASON for VCAT/LCAS has recently been undertaken [36], including discovery of VCAT and LCAS capable path termination sources and sinks and VCAT group identification. In the context of path-independent call component routing, the call path calculation gets a new dimension. In general, a call, i.e., call components, traverse two UNI interfaces, and one or several administrative domains (and therefore several E-NNI interfaces). Distributed call signaling and call path computation are highly intertwined: how and where a call path is calculated depends on the organization of the routing process and the corresponding availability of routing information. As already mentioned, ASON routing assumes hierarchically organized routing areas (RA) which feature hierarchical advertisement of information between different layers of aggregation (feeding-up and down) and thus support hierarchical inter-domain routing. Routing controllers within a routing area at the particular level of hierarchy build a topology of abstracted resources, links, and nodes, and therefore they provide path calculation services to call controllers according to a hierarchical, source-based, and step-by-step routing paradigm. The setting up of a call over several domains, several call segments, as well as the error recovery by means of crank-back must be carefully orchestrated. The GMPLS notify message (RSVP-TE) is used to report failures related to LSPs to the ingress node or some other nodes responsible for error recovery. The node resolving a failure may, in the future, also perform crank-back. Modification of a connection, the flexible association of calls and connections, as well as the crank-back capability, are advanced features which are still not fully covered within the GMPLS standard framework [37].

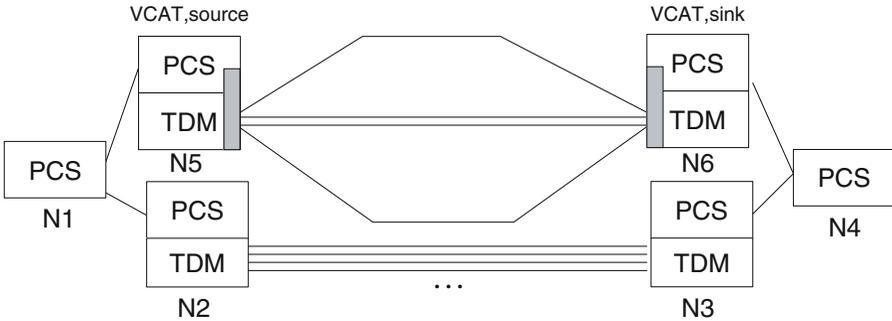


Fig. 3.4. Path-dependent vs. path-independent routing of call components

3.7 Virtual Networks

In terms of services, grooming network technology can efficiently support Layer 1 network virtualization, which is recognized as an important new operational paradigm, in particular within the GRID community [38]. Therefore, the *CP-based* support for network provision and operation of L1 virtual private network services (L1 VPN) [39] would be of very high value. Control plane support should enable L1 VPN services to use static and dynamic grooming methods to exploit the topology flexibility of grooming networks. For example, with L1 VPN, multiple virtual networks can be established over the shared infrastructure, each of which can be *selectively offered*, by means of advertisements, specific grooming and switching capabilities to establish its topology, adapting to specific traffic and connectivity requirements of applications using them. The operational benefits as compared to statically provisioned L1 VPN could be achieved with the dynamic topology engineering adapting to traffic changes, which allows resource sharing, lower cost for the customer, and better network utilization for the provider. The support currently offered with GMPLS/ASON only partially covers the functionality these advanced and challenging services would need; however, new extensions can be expected to follow to address new requirements. A related important question is whether CP support for virtualized networks would need some protocols other than GMPLS to orchestrate distributed topology engineering, or this could be done by means of signaling protocol extensions. This is an issue for further research.

3.8 Service Support

CP support to different types of services, i.e., calls or virtual networks, may differ, and in fact, CP functionality available to a specific type of services could be assumed as part of the service model, based on which a service level specification (SLS) can be defined. With a service level agreement (SLA), a guaranteed level of the service between the service user and the service provider can be further detailed.

The service model in traffic grooming networks with CP-support can include:

- *Traffic characteristics*, subject to monitoring and policing.
- The *transport plane* capabilities that the service can use, such as for example, granularity of virtual links, use of VCAT, use of grooming at different layers, transport plane sharing with the other VPNs, and resilience support.
- The capabilities of the *control plane* such as, the call set-up time, availability and support for resilience actions, maximal extensibility of the topology, the rate at which the topology can change, support for virtual concatenation and path-independent routing of components, bundling

policies, per-VPN control, support for membership discovery, supported control plane interconnection model, self-discovery over UNI, support for automatic triggering of topology changes, and control plane support for resource sharing between VPNs.

For example, the call set-up time and service availability can be used to classify services into a small number of classes (such as gold, silver, bronze) [40]. Service differentiation, based on the capabilities of the control plane, is demonstrated in the CHEETAH network [27] where the SONET path is used as a backup service for Internet transfer, and its selection depends on the expected delay of the path setup and the expected duration of the transfer. Service differentiation can also be achieved with a concept of resource visibility introduced for virtual network services [33] where each service and each network resource is associated with at least one group, and all services within one group share resources associated to it.

3.9 Conclusion

The control plane support for dynamic grooming decisions must meet two basic requirements. Firstly, there is a need for efficient modeling of network capability beyond the capacity of links at different switching layers. This grooming-specific TE functionality includes fixed and flexible inter-layer adaptations and VCAT/LCAS source and sink ends. Secondly, as the availability of network information at routing controllers along the call path is restricted due to the administrative or operational constraints related to the advertisement process, efficient mechanisms and architectures for distributed coordination of the path calculation, and for the resource allocation by signalling (along the calculated path) are required. The current work on GMPLS and PCE (path computation element) are addressing these requirements. Further extensions of CP support will play decisive role in a novel and challenging scenario of application-optimized network virtualization. Just as optimized selective deployment of grooming nodes can provide trade-off between bandwidth efficiency and network cost for an expected traffic in static grooming networks, selective access to traffic grooming capability achieved by means of selective advertised visibility of network resources can be exploited for dynamic creation of the optimized topologies for L1 VPN services. Dynamic virtual topology engineering requires especially effective orchestration of routing and signalling, as a new virtual topology link is first routed, then signalled, and then made available to attract traffic by topology information update. Topology sharing between services is another challenge for coordination of different control functionalities. Compared to future expectations, it could be observed that GMPLS/ASON currently provide a basic set of functionalities enabling automatic service provisioning, and still need to be extended to answer some of the requirements of traffic grooming. However, proposed standards have reached

some maturity, successfully demonstrated in experimental inter-operability and feasibility tests [41]. Furthermore, aiming at proof of new concepts, several demonstrators and test-beds have been recently established with different network architectures for traditional provider or grid-centric network setups [27, 42, 43, 44, 45]. These experimental studies shed light on a number of practical GMPLS aspects and on open gaps within the general standardization frameworks. They also provide a platform for studies on different service scenarios and new features which are essential for further improvements in the control plane.

References

1. H. Zhu et al., A novel generic graph model for traffic grooming in heterogeneous WDM mesh networks, *IEEE/ACM Transactions on Networking*, Vol. 11, No.2, Apr. 2003. pp. 285–299.
2. I. Cerutti et al., Traffic grooming in static WDM networks, *IEEE Comm. Magazine*, Vol. 43, No.1, Jan. 2005, pp. 101–107
3. E. Mannie, Ed., Generalized Multi-Protocol Label Switching (GMPLS) Architecture, RFC3945, Oct. 2004.
4. ITU-T G.8080 (2001) Architecture for the automatic switched optical networks (ASON).
5. S. Pasqualini et al., Influence of GMPLS on network providers' operational expenditures: a quantitative study, *IEEE Comm. Magazine*, Vol.43, No.7, Jul. 2005, pp. 28–38.
6. K. Kompella et al., Label Switched Paths (LSP) Hierarchy with Generalized Multi-Protocol Label Switching (GMPLS) Traffic Engineering (TE), RFC4206, Oct. 2005.
7. ITU-T G.707, "Network Node Interface for the Synchronous Digital Hierarchy (SDH)," Dec. 2003.
8. ANSI T1.105-2001, "Synchronous Optical Network (SONET) – Basic Description Including Multiplex Structure, Rates, and Formats," 2001.
9. ITU-T G.709, "Interfaces for the Optical Transport Network (OTN)," 2003.
10. K. Kompella et al., OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS), RFC4203, Oct. 2005.
11. K. Kompella et al., Intermediate System to Intermediate System (IS-IS) Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS), RFC4205, Oct. 2005.
12. L. Berger, Ed., Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions, RFC3473, Jan. 2003.
13. P. Ashwood-Smith et al., Generalized Multi-Protocol Label Switching (GMPLS) Signaling Constraint-based Routed Label Distribution Protocol (CR-LDP) Extensions, RFC3472, Jan. 2003.
14. L. Berger, Ed., Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description, RFC3471, Jan. 2003.
15. K. Kompella et al., Routing Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS), RFC4202, Oct. 2005.

16. J. Lang, Ed., Link Management Protocol (LMP), RFC4204, Oct. 2005.
17. S. Sengupta et al., Switched optical backbone for cost-effective scalable core IP networks, IEEE Comm. Magazine, Vol. 41, No.6, Jun. 2003, pp. 60–70.
18. Y. Zhu et al., End-to-End service provisioning in multigranularity multidomain optical networks, IEEE Int. Conf. Commun., Vol. 3, 20-24 June 2004, pp. 1579–1583.
19. R. Ramaswami, Optical networking technologies: what worked and what didn't, IEEE Comm. Magazine, Vol.44, No.9, Sept. 2006, pp. 132–139.
20. ITU-T G.7043, "Virtual Concatenation of Plesiochronous Digital Hierarchy (PDH) Signals," Jul. 2004.
21. ITU-T G.7042, "Link Capacity Adjustment Scheme (LCAS) for Virtual Concatenated Signals," Feb. 2004.
22. G. Bernstein et al., VCAT-LCAS in a clamshell, IEEE Comm. Magazine, Vol.44, No.5, May 2006, pp.34–36.
23. S. Tomic et al., GMPLS-based Exchange Points: Architecture and Functionality, Workshop on High Performance Switching and Routing, HPSR, Jun. 2003.
24. ITU-T G.805 (03-2000) Generic functional architecture of transport networks.
25. K. Kompella et al., Signalling Unnumbered Links in Resource ReSerVation Protocol – Traffic Engineering (RSVP-TE), RFC3477, Jan. 2003.
26. D. Papadimitriou et al., "Generalized Multi-Protocol Label Switching (GMPLS) Protocol Extensions for Multi-Region Networks (MRN)," IETF Internet Draft, Feb. 2005.
27. M. Veeraraghavan et al., On the use of connection-oriented networks to support grid computing, IEEE Comm. Magazine, Vol. 44, No.3, Mar. 2006, pp. 118–123.
28. K. Kompella et al., Link Bundling in MPLS Traffic Engineering (TE), RFC4201, Oct. 2005.
29. T. Nadeau, A. Farrel, Eds., Generalized Multiprotocol Label Switching (GMPLS) Traffic Engineering MIB, RFC4802, Feb. 2007.
30. D. Papadimitriou, OSPFv2 Routing Protocols Extensions for ASON Routing, Internet Draft, Mar. 2007.
31. J. Szigeti et al., Efficiency of information update strategies for automatically switched multi-domain optical networks, 7th International Conference Transparent Optical Networks, Vol. 1, 3–7 Jul. 2005, pp. 445–454.
32. A. Farrel et al., A Path Computation Element (PCE)-Based Architecture, RFC4655, Aug. 2006.
33. S. Tomic et al., Performance Analysis of Infrastructure Service Provision with GMPLS-based Traffic Engineering, *to appear in* Special Issue on Multi-Layer Traffic Engineering, JSAC, 2007.
34. E. Mannie et al., Generalized Multi-Protocol Label Switching (GMPLS) Extensions for Synchronous Optical Network (SONET) and Synchronous Digital Hierarchy (SDH) Control, RFC4606, Aug. 2006.
35. D. Papadimitriou, Ed., Generalized Multi-Protocol Label Switching (GMPLS) Signaling Extensions for G.709 OTN Control, RFC4328, Jan. 2006.
36. G. Bernstein et al., Operating Virtual Concatenation (VCAT) and the Link Capacity Adjustment Scheme (LCAS) with Generalized Multi-Protocol Label Switching (GMPLS), Internet Draft, Oct. 2005.
37. D. Papadimitriou, J. Drake, J. Ash, A. Farrel, L. Ong, Requirements for Generalized MPLS (GMPLS) Signaling Usage and Extensions for Automatically Switched Optical Network (ASON), RFC4139, Jul. 2005.

38. Grid Networks: Enabling Grids with Advanced Communication Technology F. Travostino (Editor), J. Mambretti (Co-Editor), G. Karmous-Edwards (Co-Editor) Wiley Publishers, Sep. 2006
39. T. Takeda et al., Layer 1 VPNs: driving forces and realization by GMPLS, *IEEE Comm. Magazine*, Vol. 43, No.7, Jul. 2005, pp. 60–67.
40. W. Fawaz et al., Service level agreement and provisioning in optical networks, *IEEE Comm. Magazine*, Vol. 42, No.1, Jan. 2004, pp. 36–43.
41. J.D. Jones et al., Interoperability update: dynamic ethernet services via intelligent optical networks, *IEEE Comm. Magazine*, Vol. 43, No. 11, Nov. 2005, pp.4-10.
42. C. Cavazzoni et al., Achievements of the EU NOBEL project, *European Conference on Optical Communication, ECOC'05*, Vol. 5, 25–29 Sep. 2005 pp. 37–40.
43. C. Pinart et al., On managing optical services in future control-plane-enabled IP/WDM networks, *Journal of Lightwave Technology*, Vol. 23, No.10, Oct. 2005, pp. 2868–2876.
44. L. Zhou et al., WDM optical network testbed and distributed storage application, *IEEE Communications Magazine*, Vol. 44, No.2, Feb. 2006, pp. 23–29.
45. D. Papadimitriou et al., Generalized Multi-Protocol Label Switching (GMPLS) Unified Control Plane Validation, *IEEE Int. Conf. Commun., ICC'06.*, Vol. 6, June 2006, pp. 2717–2724.

Grooming Mechanisms in SONET/SDH and Next-Generation SONET/SDH

Rudra Dutta, Ahmed E. Kamal, George N. Rouskas

4.1 Introduction

Synchronous Optical Network (SONET) and Synchronous Digital Hierarchy (SDH) are very closely related standards, which came into being primarily as a means of transporting telephone traffic in large volumes utilizing optical fiber transmission systems. SONET was brought forth by Bellcore (Telcordia), with coordination from International Telecommunication Union (ITU) as well as other standards organizations, and is primarily in use in North America, whereas SDH was developed by ITU, came into use slightly later than SONET, and is in use in the rest of the world. The two standards so closely resemble each other in most significant concepts as well as details that they are often spoken of as a single entity, and denoted as SONET/SDH. In what follows in the rest of this chapter as well as elsewhere in the book, the distinctions between the two are not important, and we continue to mean “SONET/SDH”, even when we mention just one of them, for ease of reference.

From the grooming point of view, SONET/SDH are very important entities. Transmission systems for telephony, originally analog, adopted digital systems starting with the introduction of the T-carrier system in 1961. From the very beginning, such systems adopted a “base rate” which was defined by the digitization of a single voice line, and successively higher rates formed by multiplexing larger number of lower rate channels, thanks to the hierarchical nature of the telephony architecture; this is the origin of the term *digital hierarchy*.

The term *Plesiochronous Digital Hierarchy* (PDH) is used to describe all such digital standards before the advent of SONET/SDH. This indicates a system in which all parts operate on clock signals that have exactly the same rate (within a bounded error), but may have different phases. Such a characteristic is arguably not a design decision as much as it is an adjustment to the realistically inevitable. The move to optical transmission systems was accompanied by a move to a *synchronous* architecture, which means that SONET/SDH performs multiplexing in a strictly time division multiplexed

manner. Input and output clocks at every network element are synchronized, so tight synchronization of the channels at all the different rates is possible across the entire network. There may be jitter and phase difference at the ingress to the SONET network, which are taken care of by introducing the appropriate amounts of buffer delay at the ingress. Once inside the network, however, the payloads and frames remain synchronized.

This approach makes it possible, for the network designer, to truly consider the issue of multiplexing of lower rate traffic streams into higher rate channels, hierarchically, as one abstracted by the network. Further, SONET introduced the concept of *concatenation* to efficiently carry higher rates than the base rate, as well as *virtual tributaries* to efficiently carry sub-rate payloads. In this sense, SONET/SDH for the first time delivered multi-rate capability as a service provided by the network, though it was not seen as such: the capability was seen entirely as being internal to the network provider, the end-user's only interface being a voice-rate device, i.e., a telephone. The seeds of grooming, however, were quietly sown.

Subsequently, SONET has undergone changes even more pertinent to grooming considerations, in its extension into Next Generation SONET (NG-SONET). The changes from PDH to SDH and into NG-SONET have provided a realistic basis on which traffic grooming approaches can be postulated. Hence our assertion that SONET and NG-SONET are important entities for the grooming network designer to be aware of. Accordingly, we briefly survey the salient features of these standards in this chapter. We refer the interested reader to more comprehensive sources such as the relevant standards [1, 2, and related], or texts on the subject such as [3].

4.2 The SONET/SDH Standard

SONET operates at four layers, as shown in Fig. 4.1. From the point of view of functionality, the lower two correspond roughly to the physical layer of the OSI reference stack, and the higher two to the data link control layer. The second highest layer also covers the add/drop functionality, which may be

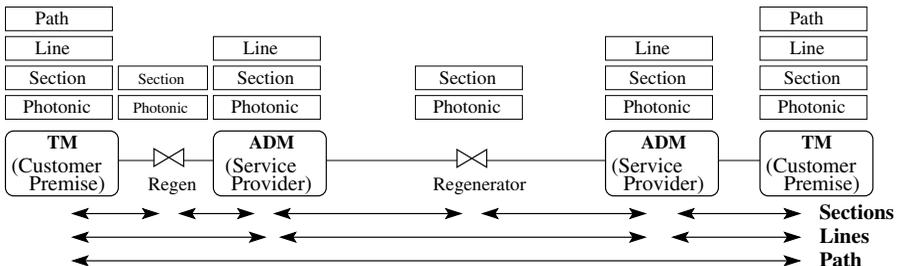


Fig. 4.1. The four layers of SONET

considered a rudimentary forwarding and therefore networking layer function, in which case the highest layer should be considered an end-to-end transport layer. However, SONET/SDH predates the OSI reference layer, and more importantly is most often used to transport frames for other technologies which define their own framing, such as ATM, Ethernet, or Frame Relay. Networking technologies adapted to such framing also tend to be second-hand users of SONET through such framing layers. Thus it is customary to look upon all four layers of SONET as making up sub-layers of the physical layer.

The Photonic layer affects the transport of bits across the physical medium, using lasers/LEDs and optical receivers, and is terminated at every physical device. The Section layer works between regenerators and repeaters in the optical transmission lines, and deals with error monitoring, signal scrambling, etc. The Line layer works at the segment between two SONET devices which understand multiplexing, and provides protection against failures. Such a segment is also called a *maintenance span*. The Path layer works between end equipments, it is a customer-to-customer transmission layer. The endpoints of the Path layer are SONET customer premise equipment, which usually is not a desktop or server, but a Terminal Multiplexer (TM), which multiplexes many end station traffic streams. The Line layer can operate between multiple ADMs or between ADMs or TMs in any combination, while the Path layer operates between TMs. The three upper layers add overhead bytes to the SONET frame.

Each level of the digital hierarchy in SONET has an associated Optical Carrier (OC) level, and an associated electrical frame structure called a Synchronous Transport Signal (STS). In SDH, a single level definition called Synchronous Transport Module (STM) is used. It is more appropriate to speak of STS frames at layers above the Photonic, and of OC signals at the Photonic layer. In this sense, the role of the Photonic layer is to map STS frames onto OC signals. Table 4.1 shows the commonly used rates, together with the corresponding bit transfer rates, both raw and without overhead. One of the distinctions of SONET/SDH from the old PDH approaches is that the fractional overhead remains constant at all levels of the digital hierarchy - in older systems they usually increase at higher levels of the hierarchy.

Table 4.1. The “digital hierarchy” of rates in SONET/SDH

SONET Level Optical/Electrical	SDH Level	Line Rate (Mbps)	Payload Rate (Mbps)
OC-1/STS-1	(STM-0)	51.840	50.112
OC-3/STS-3	STM-1	155.520	150.336
OC-12/STS-12	STM-4	622.080	601.344
OC-48/STS-48	STM-16	2488.320	2405.376
OC-192/STS-192	STM-64	9953.280	9621.504
OC-768/STS-768	STM-256	39818.120	38486.016

thus it takes the same 125 μ s to send, allowing it to keep up with every phone conversation multiplexed in it. The frame is formed by a column-wise multiplexing of the constituent STS-1 frames.

This straight mapping from STS-1 to STS-N frames creates a so-called *channelized* frame, i.e., an STS-3 frame contains three channels each of rate STS-1. SONET also allows a *non-channelized* (or unchannelized) multiplexing, through a mechanism called *concatenation*. A concatenated frame is indicated by a small “c” at the end. An STS-3c frame does not represent three STS-1 frame (with three SPEs), but a frame with a single large SPE. The Path overhead still forms the first column of the SPE, but now there is only one Path overhead in the entire STS-3c frame. The pointers in the Transport overhead have different structure for concatenated frames, and this indicates the concatenation. This allows super-rate payloads, i.e., payloads with rates higher than the base rate, to be carried by SONET/SDH.

There is also a mechanism to carry payloads with rate lower than the base rate without wasting rest of the STS-1 frame - this is called Virtual Tributaries (VT) which is the term used for the lower rate data streams. The payload columns of an STS-1 frame using VTs is divided into seven VT groups (VTGs) of 108 bytes (12 columns) each. Four types of VTs of sizes 27 bytes, 36 bytes, 54 bytes and 108 bytes are defined. Each VTG in the frame contains either four 27-byte VTs, or three 36-byte VTs, etc. VTGs are allowed to float inside the STS frame just like the SPE; there is also a locked mode which forces the VTGs to begin at the beginning of the payload columns. While not perfectly general, these capabilities provide quite a bit of flexibility in carrying lower rate traffic, and later in this chapter we see how NG-SONET extends these capabilities.

4.3 SONET Ring Networks

SONET is defined to operate as a point-to-point, linear, or ring network (in the access context, a hub configuration is also possible by combining multiple point-to-point networks). The first two are obviously special cases of the ring configuration, and the ability of SONET to operate as ring networks has been its defining identity until recently, as is the Automatic Protection Switching (APS) or the *self-healing* character of SONET rings (though APS is also defined for other configurations). In this section, we briefly describe SONET rings and APS.

SONET rings are widely used in metropolitan access and backbone, or interoffice, networks. A typical WDM SONET ring is shown in Fig. 4.3. At each node, one or more wavelengths can be dropped and/or added using an Optical Add/Drop Multiplexer (OADM). An Optical Cross Connect (OXC) can also be used for the same purpose, in addition to switching wavelengths between fibers. The dropped wavelengths are then handled by SONET Add/Drop Multiplexers (ADM) after conversion to the electronic domain

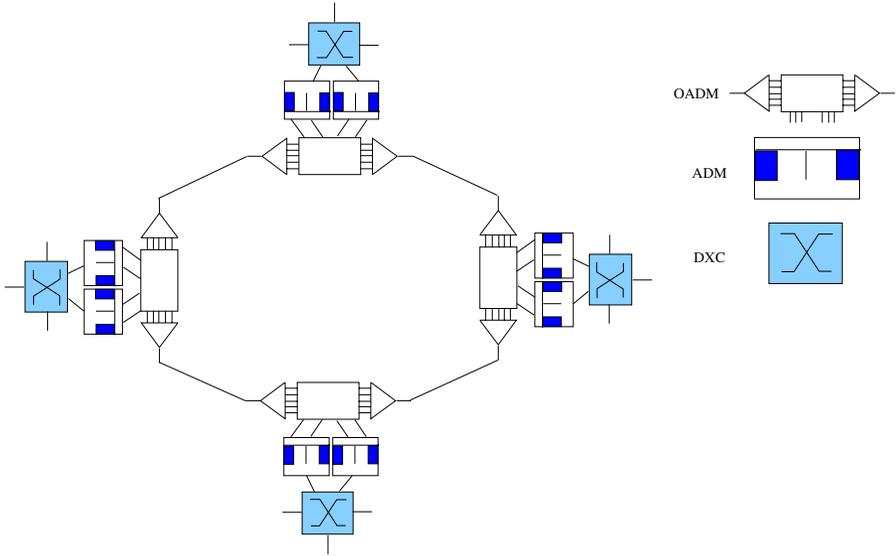


Fig. 4.3. A WDM SONET ring

using transponders. If WDM is not used, then OADMs are not needed, and the fiber is terminated at each ADM through the transponder. (This is the original mode of operation of SONET on single-wavelength systems.) The ADM is capable of adding and dropping lower speed data tributaries to and from the stream received on the ring. The ADM has a high speed interface to the ring, e.g., OC-192 for backbone rings and OC-12 for access rings. It also has a low speed interface that is typically connected to a Digital Cross Connect (DXC). The DXC is used to crossconnect lower speed streams and manage all transmission facilities in the interoffice. If multiple rings are interconnected, the DXC switches traffic between such rings and is capable of supporting a number of subtending rings. Newer technology, and advances in SONET, viz., Next Generation SONET (NGS), has led to the integration of the ADMs and the DXC functionality into one device, the Multiservice Provisioning Platform (MSPP). The MSPP has the added functionality of crossconnecting between multiple fibers.

SONET rings are usually configured in a manner that provides a survivable mode of operation in the case of equipment or fiber failure. This is one attractive feature of rings, since the ring topology is the least expensive bi-connected topology, which is the feature required to withstand single failures. Such rings are usually known as *self-healing* rings, and they are provisioned and deployed in one of two architectures, depending on the SONET sublayer at which survivability is provided:

- The Unidirectional Path Switching Ring (UPSR), and
- The Bidirectional Line Switching Ring (BLSR), which may employ either two (BLSR/2) or four (BLSR/4) fibers.

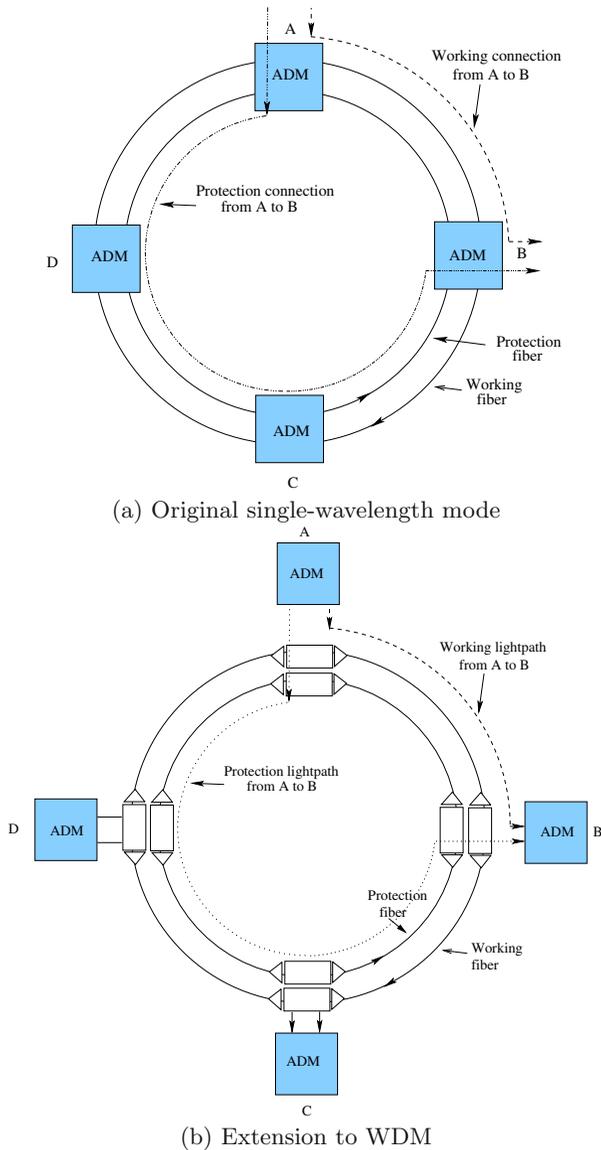


Fig. 4.4. A UPSR ring with each connection transmitting on a working and protection path

Both ring topologies guarantee that failures will be recovered from within the industry standard of 60 ms.

The UPSR provides protection at the *path* sublayer by using two fibers for transmission in two opposite directions (see Fig. 4.4). One fiber is used as the working fiber and the other is used as the protection fiber. The information in a connection from an ADM at a node to another ADM at another node,

is transmitted on both fibers (paths) at the same time. The failure of any link that affects the working fiber can be tolerated due to the reception of the signal on the protection fiber (see Fig. 4.4). This mode of survivable operation belongs to 1+1 protection and recovery from failures is instantaneous, but the

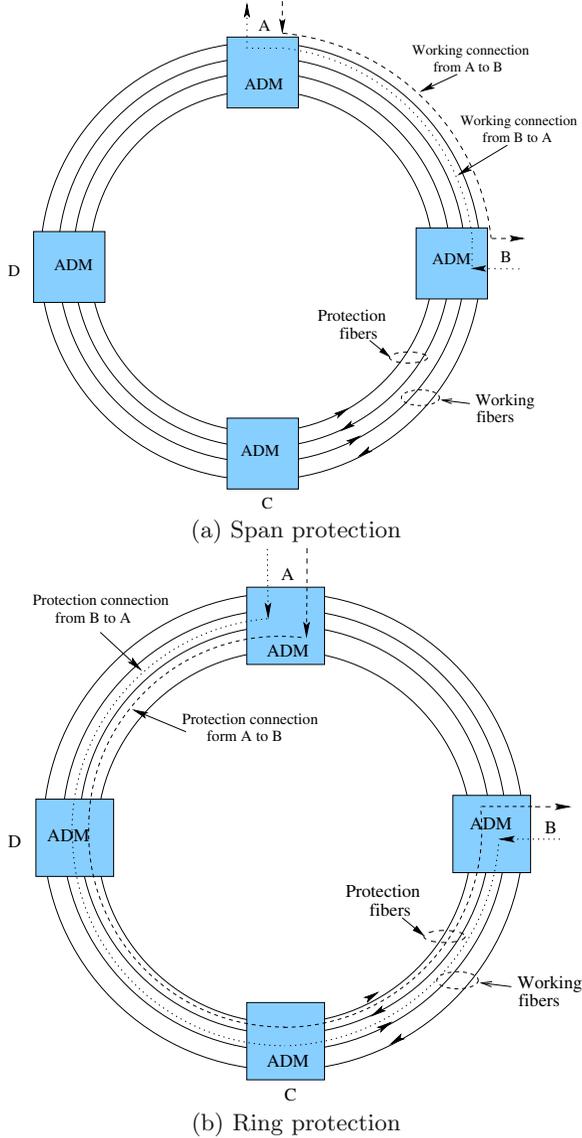


Fig. 4.5. A BLSR/4 ring: (a) Working as well as protection connections are provisioned using the shortest path routes; (b) Ring protection using the protection fiber, and at the line sublayer

ring bandwidth is not efficiently utilized since protection bandwidth is not shared between connections.

Incidentally, Fig. 4.4 also shows how these concepts apply to the WDM SONET ring in Fig. 4.3, treating each wavelength as a separate virtual link. Part (a) shows the original single-wavelength UPSR and Part (b) shows the situation for WDM rings; the key observation is that the protection path must continue to be routed through the separate protection fiber (most likely on the same wavelength as the working lightpath). Extending the BLSR SONET-based protection which we discuss next to WDM SONET rings is similarly straightforward, and we show only the original modes in the rest of this discussion.

The BLSR/4 ring uses two working fibers in two opposite directions, and similarly two protection fibers, also in two opposite directions. In BLSR/4, a connection between two nodes is provisioned using the ring with the shortest path between the two nodes (see Fig. 4.5a). If the working fibers between two nodes fail, e.g., between nodes A and B in Fig. 4.5a, then the protection fiber between the same pair of nodes, in the same direction, can be used for protection. This is known as *span switching*. However, since working and protection fibers between a pair of nodes usually share the same conduit, the failure of a span between a pair of nodes usually indicates the failure of all the fibers between this pair of nodes. Therefore, once a working fiber fails, the failure is detected by the source node, and the signal is switched to the protection fiber in the opposite direction, as shown in Fig. 4.5b. This is known as *ring protection* and is provided at the *line* sublayer. The advantage of the BLSR/4 ring is that the protection capacity is not dedicated to a specific connection, and it may even be used to carry low priority traffic, which may be preempted due to the failure of working fibers.

The BLSR/2 is similar to the BLSR/4 except that it has two fibers for communication in two opposite directions. However, the capacity of each fiber is divided into two halves: one half for working capacity and the other half is for protection capacity. Therefore, there are four virtual rings, which can be used in exactly the same way as in BLSR/4.

4.4 Next-Generation SONET/SDH (NG-SONET/SDH)

As we remarked at the beginning of this chapter, SONET/SDH networks evolved over the years to provide efficient and robust transport of voice services over long distances. Due to the characteristics of speech as well as historical and economic considerations, these networks were optimized for voice by defining a rigid hierarchy of channel capacities that are fixed multiples of 64 Kbps. SONET/SDH attributes were designed so that multiplexing operations can be performed cost-effectively by equipment of relatively low hardware complexity on such digitized voice streams.

Given the dominance, until recently, of voice traffic over data traffic, it is no surprise that the SONET/SDH standards gave little consideration to any data protocol that might be carried over these networks. Traffic trends, however, have changed significantly since the middle of the previous decade. Data traffic has overtaken voice traffic in terms of volume and continues to grow at more rapid rates. More importantly, following the Internet's ubiquity as a globally accessible data network, traffic patterns have evolved from the local concentrations of the past to traffic widely distributed over large geographical areas. As a consequence, the SONET/SDH infrastructure is increasingly used for transport of various data services, including Ethernet, Frame Relay, and Fibre Channel. These protocols were typically designed and optimized for short reach, were developed independently of optical transport networks, and did not make any attempt to leverage the capabilities of these networks.

Next-generation SONET/SDH refers to a set of standardized solutions that address the challenge of providing Data over SONET (DoS) services so as to accommodate protocols that were not developed with the transport network in mind, while allowing the flexibility to support new protocols [4]. Specifically, the enhancements include three elements:

- **Virtual concatenation (VCAT)**. This is a technique that overcomes some of the rigidities in the bandwidth hierarchy of SONET/SDH, and allows a more efficient choice of channel capacities.
- **Link capacity adjustment scheme (LCAS)**. LCAS refers to a set of procedures for adjusting dynamically the size of virtually concatenated channels.
- **Generic Framing Procedure (GFP)**. A robust yet lightweight and simple mechanism for adapting data traffic onto byte-synchronous channels, including SONET/SDH.

The new enhancements allow NG-SONET/SDH networks to improve their effectiveness in terms of grooming packet traffic, as well as their ability to accommodate demands that vary over time. The next three subsections describe each of these mechanisms in detail.

4.4.1 Virtual Concatenation

Virtual concatenation [4] generalizes the contiguous concatenation mechanism of SONET/SDH by introducing several new features that significantly enhance both the efficiency with which channels can be bundled to more closely match specific data services, and the flexibility of selecting these channels and routing them over the underlying network. With virtual concatenation, network operators can bundle N low-capacity channels to create a single channel with N times the capacity of the individual ones. The resulting high-capacity channel is referred to as the virtual concatenation group (VCG), and the individual channels in the VCG are called group members.

Two types of virtually concatenated signals have been standardized. In *high-order* virtual concatenation, N STS-1 (respectively, STS-3c) channels can be grouped to form a single STS-1- Nv (respectively, STS-3c- Nv) pipe, where “v” stands for “virtual”; in this case, the number N of lower rate signals may be any integer between 1 and 256. In *low-order* virtual concatenation, N VT1.5, VT2, VT3, or VT6 channels can be grouped to form one VT1.5- Nv , VT2- Nv , VT3- Nv , or VT6- Nv channel, respectively; here, N may vary between 1 and 64. Hence, whereas in contiguous concatenation the number N of channels to be concatenated is determined by the bandwidth hierarchy; in virtual concatenation N can be any arbitrary integer within the specified ranges.

An important feature of virtual concatenation is that the channels to be concatenated do not have to use contiguous slots in the SONET/SDH frame, nor do they have to travel over the same path. In other words, any set of N channels (say, STS-1 pipes) that originate and terminate at the same pair of path terminating equipment (PTE) can be combined into the same VCG. In fact, the virtual concatenation functionality is implemented exclusively at edge nodes. Interior network nodes treat the constituent channels of a VCG independently, as they have no way of associating them with their group; this association takes place only at the end points. Edge implementation of virtual concatenation is of practical importance as it makes it possible for network operators to gradually introduce this functionality simply by upgrading the edge nodes, without the need to introduce any modifications to the core network infrastructure. Note also that channels of a VCG that take different paths to the destination, in general, experience different delays. The destination PTE employs synchronization buffers to eliminate the differential delay and reconstruct the original data.

Figure 4.6 illustrates how a Gigabit Ethernet (GbE) signal can be transported over a SONET/SDH network using virtual concatenation. Since $N = 7$ STS-3c channels are required to carry the GbE signal, the virtual concatenation module at the originating PTE combines 7 individual STS-3c channels, all terminating at the destination PTE, into a single STS-3c-7v pipe. The 7 STS-3c channels need not occupy contiguous slots, and in fact they may travel

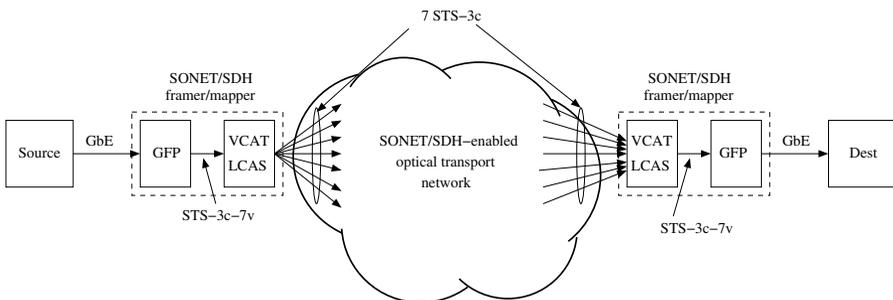


Fig. 4.6. Virtual concatenation for transport of Gigabit Ethernet (GbE)

over different paths to the destination, as shown in the figure. The GbE signal at the source is first mapped onto the STS-3c-7v VCG, e.g., using the generic framing procedure discussed shortly, and the data is carried to the destination node along the paths taken by the various group members. Once at the terminating PTE, the virtual concatenation module assembles the incoming group members into the STS-3c-7v channel, after adjusting for the delay differences, which in turn is mapped back to a GbE signal.

Virtual concatenation provides a much finer granularity in allocating bandwidth to client data signals, resulting in significant bandwidth savings compared to contiguous concatenation. Consider, for example, carrying a 100 Mbps Fast Ethernet signal over a SONET/SDH network. With contiguous concatenation, it is necessary to round the bandwidth demand to the nearest applicable SONET/SDH signal. Hence, the Fast Ethernet source must be mapped onto an STS-3c channel at 155 Mbps, an inefficient solution that wastes one-third of the allocated bandwidth. With virtual concatenation, on the other hand, the Fast Ethernet source is mapped onto a VT1.5-64v VCG, a solution with a bandwidth efficiency of 98%. Because of its finer granularity, the rounding error is much smaller with virtual concatenation, resulting in higher bandwidth utilization. Table 4.2 provides the efficiency of carrying some common data protocols over SONET/SDH networks with and without concatenation.

Another benefit of virtual concatenation is in reducing the fragmentation of spare bandwidth. With contiguous concatenation, concatenated signals must take the same path to the destination and must be on adjacent SPEs (slots). These requirements may lead to fragmentation in both time (i.e., when sufficient capacity exists in a frame but is not contiguous) and space (i.e., when sufficient capacity exists between source and destination but is distributed over different network paths). Virtual concatenation overcomes this problem as it makes it possible to use any available capacity by grouping together non-contiguous channels and/or channels over different paths to form a VCG.

Table 4.2. Bandwidth efficiency of virtual concatenation

Data signal	SONET/SDH payload (efficiency)	SONET/SDH with VCAT (efficiency)
Ethernet	STS-1 (21%)	VT1.5-7v (89%)
Fast Ethernet	STS-3c (67%)	VT1.5-64v (98%)
ESCON	STS-12c (33%)	STS-1-4v (100%)
GbE/Fibre Channel	STS-48c (40%)	STC-3c-7v / STS-1-21v (95% / 98%)

Finally, virtual concatenation provides a way to partition SONET/SDH bandwidth into several sub-rates, each of which may accommodate a different service, thus allowing multiple distinct client data signals to share, and co-exist onto, the same SONET/SDH OC-n channel.

4.4.2 Link Capacity Adjustment Scheme

The link capacity adjustment scheme (LCAS) protocol [5] is a more recent enhancement to virtual concatenation that makes it possible to increase or decrease dynamically the capacity of a VCG by adding or removing, respectively, members of the VCG. LCAS is triggered at the source node of a VCG, which exchanges signaling messages with the remote end to synchronize the addition or removal of SONET/SDH channels from the VCG. Such adjustment may be made in response to a network failure that affects one or more group members, or to time-varying traffic demands.

Dynamic Bandwidth Allocation for Time-Varying Demands

LCAS allows carriers to assign and reallocate bandwidth on the fly so as to accommodate traffic demands that change over time, and hence increase the utilization of their network. One practical application of LCAS is in adjusting the bandwidth along certain routes on a time-of-day basis, whenever traffic variability is predictable and seasonal, or even to accommodate traffic burstiness. In this case, the network management system or the call admission control system would monitor the bandwidth requirements of each VCG, and issue explicit instructions to trigger LCAS. Since it is important to ensure that any adjustment to capacity is performed in a “hitless” manner (i.e., without any data loss or bit errors during the process), the source and destination nodes employ a handshake protocol. For example, after a request to increase capacity, a new channel is provisioned and established; only after the new group member is verified and acknowledged does the source begin to send data over it. A similar process takes place when it is necessary to decrease capacity. The signaling information is exchanged in the H4 byte in the path overhead of the SONET/SDH frame, thus ensuring hardware-level synchronization.

Soft Failures for Data Traffic

As we discussed in the previous section, the ring protection mechanisms were designed to redirect all the channels carried by a failed link to a diversely routed backup path. These mechanisms are consistent with the original design objective of SONET/SDH technology, namely, as infrastructure for transporting voice calls. Since voice calls are either carried in their entirety or blocked, in the context of voice traffic SONET/SDH channels have only a binary status: either working correctly or failed. However, when the network carries data

traffic, especially elastic traffic regulated by TCP's congestion control mechanism, the status of a link may take a range of values from less to more congested, e.g., as determined by the fraction of dropped packets. In other words, a link that experiences a drop in capacity, say, from 1 Gbps to 850 Mbps, remains available for carrying data traffic. This drop in capacity is referred to as a "soft" failure, in contrast to a "hard" failure that causes the link capacity to be lost in its entirety. The LCAS failure mechanism can be used to provision links that exhibit these soft failure characteristics appropriate for a wide spectrum of data services.

Recall that virtual concatenation allows a single client signal (e.g., Gigabit Ethernet) to be carried over a VCG whose group members may take different paths across the network. Let us also make the assumption that the network operator has the capability to provision VCG members over paths that are diversely routed across the network. If one of the group members fails (e.g., due to a link cut along its path), the LCAS failure mechanism is triggered and the size of the VCG is reduced to the number of surviving members, i.e, those unaffected by the failure. As a result, the client data service may continue to use the VCG, albeit at a reduced capacity. Similarly, once the failure has been restored, the size of the VCG can be increased accordingly.

4.4.3 Generic Framing Procedure

The proliferation of IP, Ethernet, and storage area network (SAN) technologies during the 1990s led naturally to the need to carry various types of data traffic over the existing SONET/SDH infrastructure. Transporting data over byte-synchronous SONET/SDH channels requires an *adaptation* mechanism to map the data from its native form onto the SONET/SDH format at the source, and perform the inverse mapping at the destination to reconstruct the original data from the TDM signal. For instance, it is important to have a method for delineating the boundaries between the packets of a data stream, as well as filling the gaps between successive packets with empty bits that can be recognized as such and discarded at the destination.

A variety of adaptation mechanisms have been developed to map data signals over transport networks [6]. Packet over SONET (POS) is a standardized solution for mapping IP packets into SONET/SDH frames [7]. In this approach, IP datagrams are encapsulated into Point-to-Point protocol (PPP) packets, which are then framed using High-level Data Link Control (HDLC). In other words, PPP performs the mapping and encapsulation of data, while HDLC provides for delineation (or demarcation) of the PPP packets using a special flag byte. Concurrently with these standardization efforts, a number of proprietary adaptation mechanisms were developed for data over SONET/SDH mappings, creating major obstacles for interworking between equipment from various vendors. Hence, towards the end of the last decade, it was clear that a standardized solution was needed.

Generic framing procedure (GFP) [8, 9] provides a standard [10] and lightweight mechanism for mapping a variety of data signals onto a synchronous transport stream, including SONET/SDH frames, the optical transport network (OTN), or point-to-point fiber links. Figure 4.7 illustrates the relationship of GFP to client signals and underlying transport network, while Fig. 4.8 shows the structure of a GFP frame. GFP frames consist of a 4-byte core header followed by a variable-length payload with a maximum size of 64 KB. The first two bytes of the core header specify the payload length, while the last two bytes (header error control) are a cyclic redundancy check that protects against errors in the core header.

GFP functionality consists of both common and client-dependent aspects, as shown in Fig. 4.7. The common aspects of GFP apply to all adapted traffic, and include two main functions:

- **Frame delineation.** The frame delineation process of GFP is shown in Figure 4.9. Under normal operation, the receiver is at the “Sync” state,

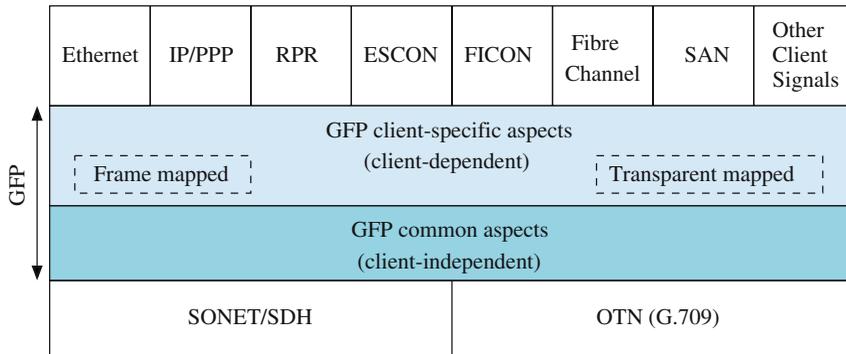


Fig. 4.7. GFP functionality

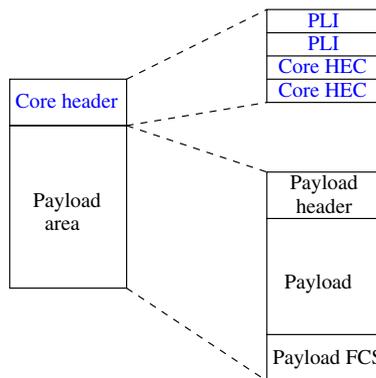


Fig. 4.8. GFP frame structure – PLI: payload length indicator, HEC: header error control, FCS: frame check sequence

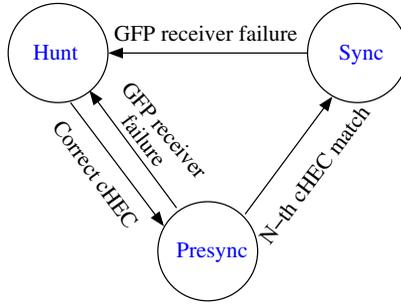


Fig. 4.9. GFP frame delineation

and it simply examines the payload length indicator in the core header to determine where the current frame ends and the next one begins. However, during link initialization or after the loss of a frame due to errors, the receiver enters the “Hunt” state hunting for a core header. This is accomplished by reading in four bytes at a time and checking the correctness of the header error control. If it is correct, the receiver transitions to the “Presync” state; otherwise the procedure is repeated. The receiver remains in the “Presync” state until a number N of frames have been correctly identified, at which time it transitions to the “Sync” state and normal operation.

- **Client multiplexing.** This feature of GFP allows several client signals of different types to share a single transport link. The multiplexing function relies on extension headers inside the payload area of the GFP frame; these headers include fields that identify the frame as belonging to a particular channel. Different extension headers are used depending on whether the multiplexing is on a single link (linear extension header) or on a ring network (ring extension header).

The client-dependent aspects of GFP perform signal adaptation (or payload mapping). Two adaptation modes have been defined. With frame mapping, each client frame (e.g., packet) is mapped in its entirety into a single GFP frame. This adaptation mode is applicable to packet-based streams, and mappings have been defined for Ethernet and IP/PPP payloads. The second mechanism, transparent mapping, is useful for delay sensitive applications, such as SAN traffic transported over Fibre Channel that requires low latency. Rather than waiting for the entire frame to be received, this adaptation mode is designed to map individual characters (code words) as they are received into GFP frames. Only client signals using 8B/10B encoding (which maps 8 bit characters to 10 bit words) may use transparent mapping. To further reduce the latency introduced by GFP mapping, the GFP frames in this case have fixed length.

4.5 Conclusion

We can thus trace the evolution of SONET from a technology crafted to telephony requirements, improving upon previous digital hierarchies by perfecting the abstraction of rate management, to a highly flexible technology that keeps its original strengths, and is friendly toward data transmission, arbitrary and dynamic rate management services as seen by the end user, and the ability to carry a wide variety of payload types. Many topics dealt with in the latter two parts of this book depend upon the existence of these capabilities, whether provided by SONET/SDH or any other technology. Early in Part II, we present studies of grooming techniques developed specifically with SONET in mind. The grooming of dynamic traffic assumes capabilities similar to LCAS. Most grooming researchers assume, either explicitly or implicitly, capabilities that parallel VCAT and GFP. As we mentioned above, the development of these capabilities have also been informed somewhat by the requirement of grooming in practice. It is likely that in the future, this synergy will continue, grooming studies and technology development continuing to inform each other.

References

1. Telcordia Standard GR-253. Synchronous optical network (SONET) transport systems: Common generic criteria. 2005. Issue 4.
2. ITU-T G.707. Network node interface for the synchronous digital hierarchy (SDH). 2003. (most current at time of writing).
3. W.J. Goralski. *SONET*. Mc-Graw Hill, 2000.
4. D. Cavendish, K. Murakami, S-H. Yun, O. Matsuda, and M. Nishihara. New transport services for next-generation SONET/SDH systems. *IEEE Communications Magazine*, 40(5):80–87, May 2002.
5. ITU-T G.7042. Link capacity adjustment scheme (LCAS) for virtually concatenated signals. 2004.
6. P. Bonenfant and A. Rodriguez-Moral. Framing techniques for IP over fiber. *IEEE Network Magazine*, 40(4):12–18, July/August 2001.
7. A. Malis and W. Simpson. PPP over SONET/SDH. IEEE RFC 1577, June 1999.
8. E. Hernandez-Valencia, M. Scholten, and Z. Zhu. The generic framing procedure (GFP): An overview. *IEEE Communications Magazine*, 40(5):63–71, May 2002.
9. P. Bonenfant and A. Rodriguez-Moral. Generic framing procedure (GFP): The catalyst for efficient data over transport. *IEEE Communications Magazine*, 40(5):72–79, May 2002.
10. ITU-T G.7041. Generic framing procedure (GFP). 2001.

Computational Complexity

Shu Huang

5.1 Introduction

The traffic grooming problem in optical networks can be posed as an optimization problem: given the constraints of the optical devices and network equipment, sub-wavelength traffic demands are aggregated onto lightpaths such that some objectives can be minimized. The traffic grooming problem has been of practical interest because of its importance in reducing the network cost, as well as research interest because of its complexity. The complexity of the traffic grooming problem partly comes from different variations of its definitions. To enumerate some, the objective of the problem may take different forms, focusing on different aspects of the network cost. The network to be studied can form a path, a ring, a tree, or mesh topology. Depending on the availability of wavelength converters, the wavelength continuity constraint may or may not be imposed. The number of wavelengths may be limited. The traffic models can be different. Given different forms of the problem, it turns out that most of them fall into the class of hard problems in terms of the computation complexity, except for some very simple cases. In this chapter, we present a detailed discussion of the complexity results in the context of the traffic grooming problem.

We first distinguish two versions of the traffic grooming problem. The static traffic grooming problem assumes that the traffic demands from users form a static traffic matrix, i.e., each traffic component of the traffic matrix, which represents the traffic demand from source node i to destination node j , is a constant. On the other hand, in the dynamic traffic grooming problem, traffic demands are variables that are functions of time. As a result, the solution may also be some functions of time. Generally speaking, the dynamic traffic grooming problem focuses on providing on-line heuristics instead of solving an Integer Linear Programming (ILP) problem optimally. For such cases, computational complexity for the entire set of traffic demands is not relevant in the same sense as for the static case. The quantity of interest is the runtime complexity of the on-line algorithms that are executed every time

a connection request arrives; this is usually low. However, in a special version of the problem where the traffic demands change only at some discrete time epochs and the changes are known beforehand, the ILP formulation can fold out to form a static problem with a bigger input size. In this case, all the complexity results from the previous sessions apply. In this chapter, we concentrate on the complexity of the static traffic grooming problem.

It is well established that the static traffic grooming problem can be decomposed into some sub-problems. At the heart of the traffic grooming problem is the virtual topology design problem, which aims at finding a graph formed by lightpaths that can carry traffic in the form of optical signals. Given a virtual topology, the traffic grooming sub-problem deals with routing traffic demands using the lightpaths. The virtual topology is mapped onto the physical topology by solving the routing and wavelength assignment problem (generally referred as RWA problem), which decides the wavelength and physical route for each lightpath. We expect that decomposing the static traffic grooming problem into sub-problems will provide us a better understanding of the whole problem as well as some “smaller” problems that may be easier to solve. However, it turns out that the second part of the expectation has not been quite successful. For one thing, all the sub-problems are tightly coupled so that solving them independently does not help us a lot in solving the whole problem. For another, the traffic grooming problem appears to be essentially hard in the sense that all the sub-problems are individually hard to solve. Indeed, although the hardness of these sub-problems can be proved by reductions from various known NP-Hard problems, they seem to share some difficulties (e.g., the integer requirements) in common. Due to these reasons, we focus on presenting the complexity of the whole traffic grooming problem.

5.2 Static Traffic Grooming

As mentioned above, the traffic grooming problem is indeed an optimization problem. For a mathematical formulation, please refer to [1] and many other papers. Notice that, to be able to examine the complexity of the problem, by convention, we need to consider its decision version. A vast majority of current deployed transmission systems are based on a physical topology of rings for its simplicity and ability of self-healing. Consequently, most early work on traffic grooming consider a ring topology. However, this simple physical topology does not make the problem significantly easier.

We start with Chiu and Modiano’s result, which, to the best of our knowledge, is the first complexity proof on the traffic grooming problem. In [2], Chiu and Modiano show that the static traffic grooming problem is NP-Complete by considering a simple case where all traffic demands are destined to one egress node in a unidirectional ring. Since each traffic demand requires exactly one ADM (Add-Drop Multiplexer) at the source node to add the traffic (without loss of generality, the traffic demand can be assumed to be less than the band-

width of the lightpath), the problem is reduced to minimizing the number of ADMs at the egress node, which is exactly the Bin Packing problem and is known to be NP-Complete. That is, this specific static traffic grooming problem can be solved in polynomial time if and only if the Bin Packing problem can be solved in polynomial time.

The proof by Chiu and Modiano is straightforward, but some assumptions are worth mentioning. Firstly, the traffic demand from one node does not bifurcate onto more than SONET rings because if it did, at least two ADMs would be required at the source node. However, by the assumption that the traffic demand is less than the bandwidth of a lightpath, we can always come up with a solution with an additional lightpath to accommodate the traffic, which requires exactly two ADMs, one at each end. Therefore, with a limited number of wavelengths, since some traffic demands may have to be split, the reduction from the Bin Packing problem seems not to be straightforward. However, as Chow and Lin show in [3], even with traffic bifurcations allowed, the reduction from the Bin Packing problem is still valid. Secondly, since the ring under study is unidirectional, the routing of traffic, an important sub-problem of the static traffic grooming problem, is left out. Indeed, this proof shows that the static traffic grooming problem in ring networks is itself an inherently difficult problem. Nevertheless, the bidirectional rings that Chow and Lin study take the routing into consideration. Therefore, the general ring grooming problem, which is to minimize the number of ADMs on bidirectional rings with traffic bifurcations allowed, is proved to be NP-Complete.

In [4], the authors demonstrate the fact that the goals of minimizing the number of ADMs and the number of wavelengths may not be possible to achieve simultaneously. Here, a little bit of retrospect of optical network study may be helpful. In the context of virtual topology design and wavelength routing and assignment, the minimization of the number of wavelengths had been extensively studied. It turns out that in ring networks, minimizing the number of wavelengths is also NP-Complete. The reduction is from the path coloring problem, which is proved to be NP-Complete in [5]. However, in [4], the authors argued that the number of ADMs reflects the network cost more accurately than the number of wavelengths. In [6], Liu et al. prove the wavelength assignment problem aiming at minimizing the number of ADMs in unidirectional rings is NP-Complete by a reduction from the circular arc coloring problem. Beyond that, in [7] they show that a different version where the rings are bidirectional is also NP-Complete. However, in these studies, researchers assume that the traffic demands are in integer number of wavelengths. As a result, when sub-wavelength traffic is taken into consideration, the hardness results need to be revisited.

In addition to the number of ADMs and wavelengths, some other objectives have also have been proposed. We enumerate them as follows:

1. *The number of Line Terminal Equipments (LTEs)*. As a variant of minimizing the number of ADMs, the problem of minimizing the number of

lightpaths is also studied. This approach is in fact integrated with the virtual topology design problem where the objective function is to minimize the cardinality of the adjacency matrix of the virtual topology, i.e., the number of lightpaths to form the virtual topology. Since each lightpath requires exactly two LTEs, this objective is equivalent to minimizing the number of LTEs. However, to take sub-wavelength traffic grooming into consideration, the number of ADMs, which consist of LTEs for adding traffic and LTEs for dropping traffic, may reflect the total cost more accurately. Again, it is shown in [4] that, subject to the same number of lightpaths, the number of ADMs of feasible solutions can be very different.

2. *The overall electronic routing involved (OEO cost)* [1]. By electronic routing, it means traffic is electronically dropped and added at an intermediate node (a node that is neither the source nor the destination). Instead of describing the cost of transceivers directly, this model focuses on the cost of electronic processing including the cost of the local tributary interfaces as well as that of switch fabric. This model provides a finer granularity to capture the overall cost of the network. In addition, from the quality of service point of view (QOS), it provides some kind of real measurement of traffic flows presented in upper layers.
3. *The network throughput or minimum blocking rate* [8]. The maximum network throughput problem and minimum network blocking rate problem are dual problems which have been studied extensively in the data network arenas. However, in the context of traffic grooming in optical networks, this model was initially ignored because traffic demands are expected to be aggregate high-speed connections in the backbone network, where it is reasonable to assume that the fluctuation caused by individual traffic demands arriving or leaving are “smoothed out”, and the aggregate traffic is relatively static. In that case it makes more sense for service providers to optimize the utilization and/or increase the capacity of the network rather than blocking them. However, more recently, attention has been paid to it.
4. *The maximum number of lightpaths originating/terminating at a transparent node.* This model is in fact a min-max approach which, instead of considering the total number of lightpaths, tries to relieve the congestion of the network. This approach does not necessarily lead to a global minimum number of lightpaths; instead the electronic routing load is more fairly distributed among the nodes. This model has been studied extensively in the context of traffic flow problems. From service providers’ point of view, this approach may be more practical than considering the number of lightpaths as a whole, because it bounds the ratio of the amount of the electronically routing traffic to the amount of the pass-through traffic at each node. On the other hand, networks with hub nodes (a node that terminates and retransmits connections between all source and destination pairs) are of interest for the flexibility and low cost in terms of the number of ADMs. In [2], it was shown that any traffic grooming scheme

that does not use hub nodes can be transformed into one that uses a hub node without adding any ADMs. However, the importance of this objective remains if distributing the switching capabilities among the network nodes is desirable.

5.3 Elemental Network Topologies

In this section, we present some complexity results of the traffic grooming problem with an objective to minimize the amount of electronic routing in path networks. The main results are adapted from [9], in which we prove that static traffic grooming problems on unidirectional path networks are NP-Complete. Propositions 1 and 2 settle the question of whether traffic grooming is tractable in a network topology as simple as the path (whether traffic bifurcations are allowed). Further, Corollary 2 shows that, for the interesting case when traffic is allowed to be bifurcated to base rate components, an approximation algorithm also cannot be hoped for. To the best of the author's knowledge, this is the first result on the approximability of the static traffic grooming problem. Notice that this result is only valid for general traffic models. An interesting result shown in [3] is a constant factor approximation algorithm for the ring grooming problem, where the traffic is uniform and the objective is to minimize the number of ADMs. In addition, the authors extend the algorithm to show that α -approximation exists for ring grooming with K -quasi-uniform traffic (i.e., the maximum traffic demand between any two nodes is bounded by k times the minimum demand between any two nodes), where $\alpha = \max(2K, 12\sqrt{2K})$.

In Section 5.3.2, we explore the consequences of these results for some other topologies of importance. In particular, Propositions 4 and 5 settle the question of the inherent difficulty of the grooming problem (as opposed to the wavelength assignment subproblem) in ring networks when the total electronic routing cost metric is used. First, we introduce some general notation and terminology.

We adopt a mathematical modeling of the grooming problem following the one given in [1]. For ease of reference, we reproduce some of the essential notation here. In the description of the traffic grooming problem, a traffic matrix $T = [t^{(sd)}]$, where $t^{(sd)}$ is the amount of traffic between source node s and destination node d , is given. It is assumed that each traffic demand is in units of a basic traffic rate (e.g. OC-3), which is not allowed to be bifurcated (routed through different lightpaths), and the number of connection may be larger than the grooming factor C , the capacity of a lightpath (e.g. OC-48) divided by the traffic rate of a connection, which is also given as part of the problem instance. We call such a $t^{(sd)}$ a traffic component or demand. We define the load L as the maximum amount of traffic traversing any physical link. Each fiber can support a given number W of wavelength channels. The network physical topology is given as a graph $G(V, E)$, with $N = |V|$.

In the following sections, the objective functions are to implicitly minimize the OEO cost if they are not explicitly defined. In practical terms, the problem of interest is the optimization problem of finding the grooming solution such that the total electronic routing over all network nodes is minimized. Instead of considering the optimization problems, we consider the simple decision versions of them. That is, given an instance of traffic grooming, the problem is whether there is a feasible traffic grooming solution such that the total amount of electronic routing is less than or equal to a given number R . Clearly, the simple decision problems are no harder than the corresponding optimization problems, since an answer to the optimization problem translates in a straightforward manner to an answer for the decision problem. On the other hand, if the decision problems are in P, we can simply use a binary search algorithm (note that the amount of electronic routing of any given instance is bounded by $\|T\|O(N)$, where T is the traffic matrix given, N is the number of nodes in the path network, and $\|T\| = \sum_{i=1}^N \sum_{j=1}^N t^{(sd)}$) is to find the minimum amount of electronic routing (a Turing reduction). Thus, the optimization problems are NP-easy[10].

5.3.1 Unidirectional Path Networks

Proposition 1. *The decision version of the traffic grooming problem in unidirectional path networks (bifurcated routing of traffic not allowed) is NP-Complete.*

An instance of the problem is provided by specifying a number N of nodes forming a network with directed connections from each node i to $i + 1$, $i \in \{1, 2, \dots, N - 1\}$, a traffic matrix $T = [t^{(sd)}]$, a grooming factor C , and a number of wavelengths supported by each link W as described in [1], and a goal R . The problem asks whether a valid virtual topology may be formed on the path and all traffic in T assigned to the virtual topology so that the total electronic routing is less than or equal to R over all the nodes. Each source–destination traffic component $t^{(sd)}$ must be routed over a unique sequence of lightpaths from node s to node d . In this context, two lightpaths which have the same source and destination nodes and follow the same physical route over the physical topology, differing only in the wavelength used, are considered the same lightpath (in keeping with [1]), and bifurcations of the same traffic component over two such lightpaths is not considered bifurcated routing. This provision is required to feasibly route traffic components which are larger than C . In such cases, the split must be to integer sub-components. It can easily be shown that requiring no more than one such sub-component to be strictly less than C does not impose a stronger constraint.

Proof. The reduction is from the Subsets Sum problem [10]. Given: an instance of the Subsets Sum problem with n elements of size $s_i \in Z^+ \forall i \in \{1, 2, \dots, n\}$, and a goal sum B . Let $B_1 = \max\{B, \sum_i s_i - B\}$. (For the purpose of the Subsets Sum problem, posing the instance with B or B_1 is equivalent.) Construct a path network using the following transformation: $N = 2n + 1$, $W = 2$,

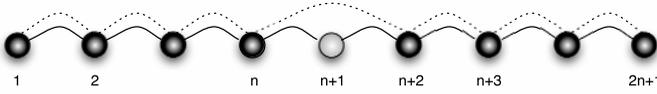


Fig. 5.1. Example of path construction of Proposition 1

$C = \sum_i s_i + 1, t^{(i,i+1)} = C + 1, \forall i \in \{1, 2, \dots, n - 1\} \cup \{n + 1, n + 2, \dots, 2n\},$
 $t^{(n,n+1)} = t^{(n+1,n+2)} = B_1 + 1, t^{(n,n+2)} = C - B_1, t^{(i,i+n+1)} = s_i, \forall i \in$
 $\{1, 2, \dots, n\},$ all other $t^{(sd)} = 0, R = n \sum_i s_i - B_1.$ An example is shown in Fig. 5.1.

Due to the traffic components of magnitude $C + 1,$ both wavelengths must be used to form single hop lightpaths over all physical links except the two central ones. Over the two central ones, at least one single-hop lightpath must be formed due to the traffic components of magnitude $B_1 + 1$ (this quantity is less than C for $0 < B < \sum_i s_i,$ i.e. non-triviality of the Subsets Sum instance, and hence can always fit in one wavelength); the other wavelength may be used to form two single-hop lightpaths over these two links, or a single two-hop lightpath over them. The electronic routing in the former case is at least as large as in the latter case, thus it suffices to consider the latter case. Thus the virtual topology may be considered forced on us by the construction. On this virtual topology, each of the traffic components corresponding to the object sizes of the Subsets Sum problem must be electronically routed exactly $n - 1$ times at all nodes other than node $n + 1.$ At node $n + 1,$ at most C units of traffic can be optically routed, since only lightpath optically passes through. The $C - B_1$ units of $t^{(n,n+2)}$ must be routed on the wavelength than bypasses node $n + 1,$ since traffic cannot be bifurcated and the other wavelength does not have enough room for it. Thus there remains room for at most B_1 units of the traffic corresponding to the object sizes of the Subsets Sum problem to optically bypass node $n + 1.$ To satisfy the electronic routing goal, at least this much traffic must be optically passed through node $n + 1,$ and because traffic cannot be bifurcated, the electronic routing goal can be satisfied *iff* there is a subset of objects of the Subsets Sum problem instance whose sizes total to $B_1,$ that is, *iff* the Subsets Sum problem instance can be satisfied. Since deciding the satisfiability of the Subsets Sum problem is NP-Complete, the proposition is proved. \square

Corollary 1. *The decision version of the traffic grooming problem in uni-directional path networks (bifurcated routing of traffic not allowed) is NP-Complete even when a candidate virtual topology is provided.* An instance of the problem is provided as above together with a valid virtual topology on the path, and the question is whether a particular value of electronic routing or lower can be achieved for the given traffic on that topology.

Proof. The proof follows the same lines as that for Proposition 1. When constructing an instance of the grooming problem, it is now unnecessary to specify

the traffic components of magnitude $C + 1$, instead, the virtual topology that was forced is specified as the candidate virtual topology. Then the grooming instance is again shown to be satisfiable *iff* the Subsets Sum instance is. \square

Note: Because of the construction in the proof, the only feasible assignment of the traffic to the virtual topology is the one that satisfies the grooming goal. Thus it is also shown that the value of R can be assigned a larger value without affecting the satisfiability of the instance. In particular, $R = n \sum_i s_i + C - B_1$ will have the same result. Since this is the maximum electronic routing that can occur for the problem instance (every traffic component is electronically routed at every intermediate node), it is also proved that the problem of deciding whether a given virtual topology admits of any feasible routing of traffic at all is also NP-Complete.

In the above proposition, we presented that the traffic grooming problem in unidirectional path networks with traffic bifurcation not allowed is NP-Complete by the reduction from the Subset Sum problem. However, it is known that the Subset Sum problem is not NP-Complete in strong sense [10]. Using a simple dynamic programming approach, a pseudo-polynomial algorithm can be employed to find an optimal solution. If the grooming goal R is bounded by a constant, which is the case in our construction, the algorithm is attractive. Unfortunately, when we eliminate the constraint that traffic components are not allowed to be routed on different lightpaths, the problem is still NP-Complete, now in the strong sense.

Proposition 2. *The decision version of the traffic grooming problem in unidirectional path networks (bifurcated routing of traffic allowed) is NP-Complete.*

An instance of the problem is provided exactly as for the case where bifurcation is not allowed, but now it is allowed to bifurcate each traffic component $t^{(sd)}$ in various subcomponents which may follow different routes from source to destination. The bifurcation is restricted to integer subcomponents.

Proof. The reduction is from the Multi-Commodity Flow (MCF) in three stage networks with three nodes in the second stage, which has been proved NP-Complete in [11]. An instance of the problem consists of a network consisting of three sets $\mathcal{N}_1, \mathcal{N}_2, \mathcal{N}_3$ of nodes forming the first, second, and third stage of a simple staged network (with $|\mathcal{N}_2| = 3$), a set of directed arcs $E \subset (\mathcal{N}_1 \times \mathcal{N}_2) \cup (\mathcal{N}_2 \times \mathcal{N}_3)$, each of unit capacity, a set of flow requirements $Q \subset (\mathcal{N}_1 \times \mathcal{N}_3)$, each of unit magnitude. The question is whether a feasible flow assignment satisfying the flow requirements exists.

We construct a path network with as many nodes as the three stage network, with a one-to-one correspondence between the nodes of the stage network and the path, as illustrated in Fig. 5.2. We define the following quantities. Let A be the set of all ordered node pairs (s, d) of the path network such that $(i, j) \in E$ for the staged network, where s is the node corresponding to i and d is the node corresponding to j . Similarly, let B be the set of all ordered node pairs (s, d) of the path network such that $(i, j) \in Q$ for the

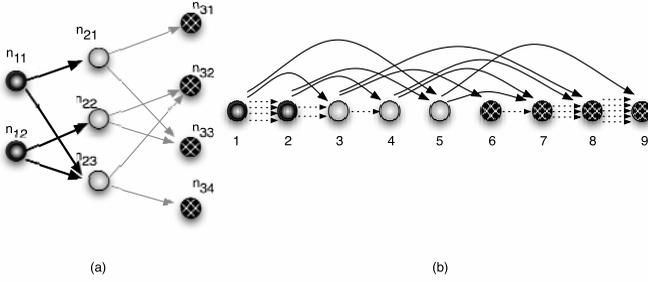


Fig. 5.2. Example of path construction of Proposition 2: (a) 3-stage network instance, (b) path instance constructed together with virtual topology forced ($W = 6$ in this instance)

staged network. For the link from node i to $i + 1$ ($i \in \{1, 2, \dots, N - 1\}$), let $w_i = |\{(s, d) : (s, d) \in A, s \leq i, d \geq i + 1\}|$. That is, w_i is the number of arcs that would cross link i if the arcs of the staged network were drawn between corresponding nodes of the path network. Construct a path network using the following transformation: $N = |\mathcal{N}_1| + 3 + |\mathcal{N}_3|$, $C = |Q| + 2$, $W = \max_i \{w_i\}$, $t^{(sd)} = C - 1$, $(s, d) \in A$, $t^{(sd)} = 1$, $(s, d) \in B$, $t^{(i, i+1)} = (W - w_i)C, \forall i \in \{1, 2, \dots, N - 1\}$. That is, each arc of the staged network generates a traffic component of magnitude $C - 1$, and each flow requirement of the staged network generates a traffic component of magnitude 1, between the corresponding nodes of the path network. All other traffic components are zero, and $R = |Q|$.

Because the magnitude of the traffic components corresponding to the arcs of the staged network are each $Q + 1$, the goal cannot be achieved if even one of these traffic components is completely electronically routed. Thus at least one unit of traffic for such a traffic component must be optically routed, and this is true for every such traffic component. Hence for the goal to be achieved the virtual topology must include at least one direct lightpath for each of these traffic components. That is, any virtual topology satisfying the goal must include at least one lightpath with source node s and d for each node pair $(s, d) \in B$. However, exactly one such, together with the $W - w_i$ single-hop lightpaths that must be formed over each link from node i to $i + 1$ to carry the single-hop traffic, will occupy every wavelength on every link; thus a complete virtual topology is forced. On this virtual topology, the single-hop lightpaths are completely occupied with the single-hop traffic. Each of the lightpaths $(s, d) \in B$ must carry the entire traffic from s to d , since there are only lightpaths from one “stage” to another in the path, and so are the traffic components. The remaining bandwidth and remaining traffic components are exactly the ones corresponding to the arc capacities and flow requirements of the MCF 3-stage problem instance. Every possible path for every traffic component involves exactly one intermediate node from source to destination, thus if it is at all feasible to route the traffic, the electronic

routing goal will be met. Thus the path network grooming problem instance is satisfiable *iff* the MCF problem instance is. Since the MCF problem is known to be NP-Complete, so is the path grooming problem. \square

We can show that this problem is NP-Complete in the strong sense. Following the same terminology in [10], let I be the instance of the traffic grooming problem transformed from the instance of the MCF problem. The instance is described by C , W , N , and T . We take the $Length(I)$ to be $N + \lceil \log_2 W \rceil + \lceil \log_2 C \rceil + \lceil \log_2 R \rceil + \sum_{sd} \lceil \log_2 t^{(sd)} \rceil$. Because the number of wavelength and the capacity of a lightpath are constrained in practice, we can take the $Max(I)$ to be the larger of R and $\max t^{(sd)}$. From the reduction, we have $R = |Q|$, where $|Q| \leq |\mathcal{N}_1| |\mathcal{N}_3| \leq |\mathcal{N}|^2$ and $\max t^{(sd)} \leq WC$. Therefore, there exists a polynomial $p(\cdot)$ such that $Max(I) \leq p(Length(I))$. Since we have shown that this problem is NP-Complete, it follows that no pseudo-polynomial time algorithm exists for the traffic grooming problem, thus it is NP-Complete in the strong sense.

Combined with the observation that the goal R is bounded by a polynomial of n and $\max_{ij} t_{ij}$, we can rule out the possibility of a fully polynomial time approximation scheme (FPTAS) unless $P = NP$. Furthermore, since the solution values are integers, it is easy to see that, for a given instance I , no polynomial time approximation algorithm A satisfying $R_A = \frac{A(I)}{OPT(I)} < 1 + \frac{1}{R}$ exists (otherwise, it implies that $1 < A(I) - OPT(I) < OPT(I)/R$, i.e., it is a polynomial algorithm that gives a YES answer to the corresponding decision problem, which contradicts the NP-Completeness proof). Thus, we also rule out the possibility of a polynomial time approximation scheme (PTAS) unless $P = NP$. However, it might be hoped that approximation algorithms may exist for some useful approximation ratios. We show in the next corollary that this is not true.

Corollary 2. *Approximating the optimization version of the unidirectional path network grooming problem (bifurcated routing of traffic allowed) is NP-hard, unless $P=NP$.*

An instance of the problem is provided exactly as for the proposition for the decision problem of traffic grooming with bifurcation allowed. Now the problem is to find the grooming solution which produces the minimum amount of electronic routing R_o .

Proof. The reduction is from the same Multi-Commodity Flow used in the proof of Proposition 2.

Suppose that we have a polynomial time approximation algorithm which has $R_M(I) \leq \infty$, for the traffic grooming problem instance I with $OPT(I) > 0$, where $OPT(I)$ is the optimal value. (Excluding the cases where $OPT(I) = 0$ does not change the intractability of the problem, since those cases are trivially solvable.) It implies that there would be a polynomial time algorithm M satisfying $R_M(I) = \frac{M(I)}{OPT(I)} \leq K$, where $M(I)$ is the result returned by the algorithm M , for some positive integer K . Then construct of the instance I

as follows: for any given instance I^{MCF} of the MCF problem, we add $K|Q|$ dummy nodes to \mathcal{N}_2 . We name them as $D=\{D_1, \dots, D_{K|Q|}\}$. First we define the following sets: A and B are sets of node pairs exactly as before. Let H be the set of all ordered node pairs (s, d) of the path network such that either $(s \in \mathcal{N}_1$ and $d = D_1)$ or $(s = D_{K|Q|}$ and $d \in \mathcal{N}_3)$. Let L be the set of all ordered node pairs (s, d) of the path network such that $s = D_i, d = D_{i+1}, \forall i \in 1, 2, \dots, K|Q| - 1$.

Then we construct a path network with as many nodes as the three stage network, exactly as in the proof of Proposition 2, with the following additions.

For the link from node i to $i + 1$ ($i \in \{1, 2, \dots, D, \dots, N - 1\}$), let $w_i = |\{(s, d) : (s, d) \in A \cup H \cup L, s \leq i, d \geq i + 1\}|$. Construct a path network using the following transformation: $N = |\mathcal{N}_1| + 3 + K|Q| + |\mathcal{N}_3|$, $C = K|Q| + 2$, $W = \max_i \{w_i\}$,

$$t^{(sd)} = \begin{cases} C - 1, (s, d) \in A \cup H \\ 1, (s, d) \in B \\ (W - w_i)C + 2, (s, d) \in L \\ (W - w_i)C, \{(s, d) | s \in \{1, 2, \dots, N - 1, d = s + 1, (s, d) \notin L\} \end{cases}$$

All other traffic components are zero, and $R = |Q|$. Since K is independent on I , this construction is in polynomial time. Since the traffic $t^{(sd)} = 1, (s, d) \in B$, can always be routed as $\{s, D_1, \dots, D_{K|Q|}, d\}$, $M(I)$ will always return a feasible solution.

If $M(I) \leq K|Q|$, it implies that none of the traffic components of magnitude $K|Q| + 1$ is completely electronically routed. Thus at least one unit of traffic for such a traffic component must be optically routed, and this is true of every such traffic component. Hence the virtual topology must include at least one direct lightpath for each of these traffic components. That is, any virtual topology satisfying the goal must include at least one lightpath with source node s and d for each node pair $(s, d) \in A \cup H \cup L$. However, exactly one such, together with the $W - w_i$ single-hop lightpaths that must be formed over each link from node i to $i + 1$ to carry the single-hop traffic, will occupy every wavelength on every link; thus a complete virtual topology is forced.

Furthermore, if even one traffic component in B is routed through the route $\{s, D_1, \dots, D_{K|Q|}, d\}$, it will introduce exactly an amount of electronic routing $K|Q|$. Since it is reasonable to assume that $|Q| > 2$ (again we exclude only trivial cases), we have $K|Q| > 2$. Therefore the amount of electronic routing is at least $K|Q| + |Q| - 1$, which is larger than $K|Q|$. Therefore, $M(I) \leq K|Q|$ implies that no traffic component in B is routed by the route $\{s, D_1, \dots, D_{K|Q|}, d\}$.

Thus, it implies that the instance I^{MCF} is satisfiable.

If $M(I) > K|Q|$, since we have $\frac{M(I)}{OPT(I)} \leq K$, it implies that $|Q| < OPT(I)$, which implies the instance of I^{MCF} is not satisfiable.

Thus, using the algorithm M , we can solve the MCF problem in polynomial time. Obviously, if we assume that $P \neq NP$, then M cannot exist. \square

Next, we prove the parallel result to Corollary 1 for bifurcated routing paths, that even with the virtual topology is given, the decision problem is still NP-Complete.

Corollary 3. *The decision version of the traffic grooming problem in unidirectional path networks (bifurcated routing of traffic allowed) is NP-Complete even when a candidate virtual topology is provided.*

Proof. This result follows from Proposition 2 in the same way as the proof for Corollary 1 does from Proposition 1. In the reduction, we construct the path network as above, but do not include the traffic components introduced to force the virtual topology; instead, the virtual topology which was forced is specified as part of the instance. The rest of the proof follows as before.

Again, by the nature of the proof, it is also proved that the problem of deciding whether a given virtual topology admits of any feasible routing of traffic at all is also NP-Complete. \square

The above also showed that the feasibility check problem, i.e., the problem that given a virtual topology, the question is that if it admits the given traffic matrix, is NP-Complete. The feasibility problem in data networks has been studied for decades. Generally, it is formulated as a multicommodity flow problem which is known to be much harder than single-commodity flow problems. In the single-commodity flow problem, the Max-Flow-Min-Cut theorem is necessary and sufficient to prove the feasibility. However, this good property does not always hold for multicommodity flow problems. One reason is that flows on a same edge do not sum up or cancel each other if they are different commodities. Obviously, in the multicommodity case, the Max-Flow-Min-Cut theorem (the max-flow is defined as the sum of multicommodity requests with source and destination separated by the cut, which is slightly different from the single-commodity case) is still necessary. Otherwise, the traffic matrix can not be satisfied. However, it is not sufficient any more. Extensive research have been reported in literature on the graphs for which the sufficiency holds. For example, in [12], the authors proved that if the graph is planar and all the sources and destinations of multicommodities can be drawn on the boundary of a infinite region, the sufficiency holds. These observation has been used on ring networks. However, our results show that the feasibility problem is intractable in even the path network because of the complexity of the virtual topology it may bring. On the other hand, in a recent work by Iyer et al., it shows that if the set of lightpaths is specified in advance, routing to minimize electronic switching costs becomes a polynomial-time problem [13]. Another interesting result shown in [13] is that, if the capacity constraint of lightpaths is relaxed, the entire problem becomes polynomial.

5.3.2 Other Topologies

In this section, we explore some of the more pertinent consequences of the propositions proved in the last section. Specifically, we show that the more

general case of bidirectional path networks is also intractable. We also show that ring network grooming problems are also intractable independent of the wavelength assignment problem complexity. This is an important result since ring networks are of current and lasting practical importance.

For the following three propositions, because of the space limit, the proofs are omitted. Interested readers should refer to [9] for details.

Proposition 3. *The decision version of the traffic grooming problem in bidirectional path networks (in both the cases of bifurcated routing of traffic allowed and not allowed) is NP-Complete.*

Proposition 4. *The decision version of the traffic grooming problem in unidirectional ring networks (in both the cases of bifurcated routing of traffic allowed and not allowed) is NP-Complete, even when every node has full wavelength conversion capability.*

Proposition 5. *The decision version of the traffic grooming problem in bidirectional ring networks (in both the cases of bifurcated routing of traffic allowed and not allowed) is NP-Complete, even when every node has full wavelength conversion capability.*

Some other interesting network topologies may also be seen to fall in the intractable class as a result of the path network intractability. These include the *spider networks* and *Manhattan street* or *grid networks*. Since these follow trivially from the path network results, we do not formally state them, nor include the demonstrations.

Next, we show that the static traffic grooming problem is NP-Complete in star and tree networks.

Proposition 6. *The decision version of the traffic grooming problem in star networks is NP-Complete.* Recalling that the problem consists of deciding which traffic elements are routed optically and which are not, we recognize that requiring the electronic switching at the hub to be F or less is equivalent to requiring that the optical routing at the hub be Q or more, where $Q = \sum_{i,j=1}^N t_{ij} - F$. In what follows, we use Q rather than F for notational convenience.

Proof. We reduce the decision version of the Knapsack problem [10] to the grooming problem. An instance of the Knapsack problem is given by a finite set U of cardinality n , for each element $u_i \in U$ a weight $w_i \in \mathbb{Z}^+$, and a value $v_i \in \mathbb{Z}^+, \forall i \in \{1, 2, \dots, n\}$, a target weight $B \in \mathbb{Z}^+$, and a target value $K \in \mathbb{Z}^+$. The problem asks whether there exists a binary vector $X = \{x_1, x_2, \dots, x_n\}$ such that $\sum_{i=1}^n x_i w_i \leq B$ and $\sum_{i=1}^n x_i v_i \geq K$. Given such an instance, we construct a star network using the following transformation: $N = n + 2$, $W = n$, $C = \max_i(w_i + v_i) + 1$, $Q = K + \sum_{i=1}^n (C - w_i - v_i)$, and traffic matrix:

$$t_{ij} = \begin{cases} C - w_j, & i = n + 1, j = 1, 2, \dots, n \\ C - w_j - v_j, & i = n + 2, j = 1, 2, \dots, n \\ (n - 2)C + w_j, & i = 0, j = 1, 2, \dots, n, \\ \sum_{k=1}^n w_k - B, & i = n + 1, j = 0 \\ 0, & \text{otherwise} \end{cases}$$

In the resulting star network, the only traffic components switched through the hub optically or electronically are those from one of the *source nodes* $n + 1$ and $n + 2$ to one of the *destination nodes* $1, 2, \dots, n$. The amount of traffic of each such component is less than the capacity of a wavelength. There is also traffic from the hub node to each destination node, and traffic from source node $n + 1$ to the hub. Due to the traffic from the hub, any one, but not both, of the traffic components from the source nodes may be optically routed for each destination node. Not all traffic sourced by source node $n + 1$ may be optically routable, due to the traffic to the hub, which requires terminating some lightpaths at the hub. There is no such restriction for source node $n + 2$, so a lightpath may be formed from it to every destination node which does not sink a lightpath from source node $n + 1$. Therefore, we need only to consider candidate solutions in which there is a lightpath from exactly one of nodes $n + 1, n + 2$, to each node $i \in \{1, 2, \dots, n\}$ to determine the satisfiability of the instance.

Let X denote a candidate solution of the Knapsack instance. Consider the solution of the star network in which X (respectively, \bar{X}) represents the indicator vector of the lightpaths formed from node $n + 1$ (respectively, $n + 2$). Applying the transformation to the satisfiability criteria of Knapsack, we obtain:

$$\begin{aligned} & \sum_{i=1}^n x_i w_i \leq B \\ \Rightarrow & \sum_{i=1}^n x_i (C - t_{n+1,i}) \leq \sum_{i=1}^n (C - t_{n+1,i}) - t_{n+1,0} \\ \Rightarrow & \sum_{i=1}^n (\bar{x}_i t_{n+1,i}) + t_{n+1,0} \leq (n - \sum_{i=1}^n x_i) C \end{aligned} \quad (5.1)$$

$$\begin{aligned} & \sum_{i=1}^n x_i v_i \geq K \\ \Rightarrow & \sum_{i=1}^n x_i (t_{n+1,i} - t_{n+2,i}) \geq Q - \sum_{i=1}^n t_{n+2,i} \\ \Rightarrow & \sum_{i=1}^n (x_i t_{n+1,i} + \bar{x}_i t_{n+2,i}) \geq Q \end{aligned} \quad (5.2)$$

Thus, the weight constraint translates to the requirement that the lightpaths from source node $n + 1$ to the hub can carry the hub traffic as well as all traffic components which have not been given a lightpath, i.e., the logical

topology is feasible. The value criterion translates to the requirement regarding the minimum amount of optical routing. Therefore, a given vector X either satisfies both the Knapsack and the grooming instance, or fails to satisfy both. Hence, the grooming instance is satisfiable *iff* the Knapsack instance is. \square

Proposition 7. *The decision version of the traffic grooming problem in tree networks is NP-Complete, even when every interior tree node has full wavelength conversion capability.* For the sake of brevity, the proof is omitted. Again, interested readers may refer to [9] for details.

5.3.3 Static Traffic Grooming with Min–Max Objective

In this section, we discuss another variant of the static traffic grooming problem, where the objective is to minimize the maximum number of lightpaths originating from or terminating at any node [14]. As mentioned in Section 5.2, this objective is of high practical value because all the network nodes are likely to be provisioned with identical equipment. It turns out that such a homogeneous network may be easier to operate, manage, and maintain, especially when the traffic model is dynamic and not known a priori. For the sake of brevity, we enumerate the main results of the paper. These results parallel very closely those obtained for the OEO minimization problem. The interested readers should refer to the paper for detailed proofs.

1. The decision version of the static traffic grooming problem in unidirectional paths with the Min–Max objective (bifurcated routing of traffic not allowed) is NP-Complete. This statement is true even when a logical topology is provided.
2. The decision version of the static traffic grooming problem in unidirectional paths with the Min–Max objective (bifurcated routing of traffic allowed) is NP-Complete. This statement is true even when a logical topology is provided.
3. The decision version of the static traffic grooming problem in bidirectional paths with the Min–Max objective (bifurcated routing of traffic allowed or not allowed) is NP-Complete.
4. The decision version of the static traffic grooming problem in unidirectional paths with the Min–Max objective (bifurcated routing of traffic allowed or not allowed) is NP-Complete, even when every node has full wavelength conversion capability.
5. The decision version of the static traffic grooming problem in bidirectional paths with the Min–Max objective (bifurcated routing of traffic allowed or not allowed) is NP-Complete, even when every node has full wavelength conversion capability.
6. If $P \neq NP$, then no polynomial time approximation algorithm A for the static traffic grooming problem in unidirectional path networks with the Min–Max objective (bifurcated routing of traffic allowed) can guarantee $A(Instance) - OPT(Instance) \leq K$ for a fixed constant integer K .

5.4 Conclusion

In this chapter, we presented some complexity results for the traffic grooming problem. We show that the traffic grooming problem has a very rich context. It takes different formulations in terms of the objectives as well as the constraints (eg., different physical topologies). The traffic grooming problem can be decomposed into some hard sub-problems that have been well studied in different context. However, using the simplest physical topologies, we show that the problem is itself a hard problem. We also show that two versions of the problem, where the total electronic processing and a Min–Max objective are minimized, are not approximable. We elaborate the proofs in detail because it builds a connection between the traffic grooming problem and a NP-hard MCF problem, which seems to be promising in solving other open questions on the complexity of the traffic grooming problem.

References

1. Dutta R. and Rouskas G.N. (2002) Traffic Grooming in WDM Networks: Past and Future, *IEEE Network*.
2. Chiu A. and Modiano E. (2000) Traffic Grooming Algorithms for Reducing Electronic Multiplexing Costs in WDM Ring Networks, *IEEE Journal on Lightwave Technology*.
3. Chow T. and Lin P. (2004) The Ring Grooming Problem, *Networks*.
4. Gerstel, O. and Lin, P. and Sasaki, G. (1999) Combined WDM and SONET Network Design, *IEEE INFOCOM*.
5. Erlebach T. (1998) Scheduling Connections in Fast Networks, Thesis, Swiss Federal Institute of Technology, Zurich.
6. Liu L.W., Li X.Y., Wan P.J., Frieder O. (2000) Wavelength Assignment in WDM Rings to Minimize SONET ADMs, *IEEE INFOCOM*.
7. Wan P.J., Calinescu G., Liu L.W., Frieder O. (1999) Grooming of Arbitrary Traffic in SONET BLSRs, *GLOBECOM*.
8. Zhu, K. and Mukherjee, B. (2002) Traffic Grooming in an Optical WDM Mesh Network, *IEEE Journal on Selected Areas in Communications*.
9. Huang S., Dutta R., and Rouskas G.N. (2006) Traffic Grooming in Path, Star, Tree Networks: Complexity, Bounds and Algorithms, *IEEE Journal on Selected Areas in Communications*.
10. Garey, M.R. and Johnson, D.S. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W.H. Freeman and Company.
11. Elmallah, E.S. and Culberson, J. (1995) Multicommodity Flows in Simple Multistage Networks, *Networks*.
12. Okamura, H. and Seymour, P.D. (1981) Multicommodity Flows in Planar Graphs, *Journal of Combinatorial Theory*.
13. Iyer P., Dutta R., and Savage C.D. (2005) On the Complexity of Path Traffic Grooming, *Proceedings of the Second International IEEE/Create-Net Workshop on Traffic Grooming*, Boston, MA.
14. Chen B.S., Rouskas G.N., Dutta R. (2005) Traffic Grooming in WDM Ring Networks to Minimize the Maximum Electronic Port Cost, *Optical Switching and Networking*.

Hierarchical Traffic Grooming

George N. Rouskas, Rudra Dutta

6.1 Introduction

Traffic grooming is the field of study that is concerned with the development of algorithms and protocols for the design, operation, and control of networks with multigranular bandwidth demands [1]. As the number of logical entities (including sub-wavelength channels, wavelengths, wavebands, and fibers) that need to be controlled in a multigranular network increases rapidly with the network size, wavelength capacity, and load, a scalable framework for managing these entities becomes a *sine qua non* for future wide area WDM networks.

Several variants of the traffic grooming problem have been studied in the literature under a range of assumptions regarding the network topology, the nature of traffic, and the optical and electronic switching model [2, 3, 4, 5, 6, 7, 8, 9, 10]. Typically, an integer linear programming (ILP) formulation serves as the basis for reasoning about and tackling the problem. Unfortunately, solving the ILP directly does not scale to instances with more than a handful of nodes, and cannot be applied to networks of practical size covering a national or international geographical area. Consequently, either the ILP is tackled using standard relaxation techniques, or the problem is decomposed into subproblems which are solved using heuristics.

Most of the above studies regard the network as a flat entity for the purposes of lightpath routing, wavelength assignment, and traffic grooming. It is well known, however, that in existing networks resources are typically managed and controlled in a hierarchical manner. The levels of the hierarchy either reflect the underlying organizational structure of the network or are designed in order to ensure scalability of the control and management functions. Accordingly, several studies have adopted a variety of hierarchical approaches to traffic grooming that, by virtue of decomposing the network, scale well and are more compatible with the manner in which networks operate in practice.

The rest of this chapter is organized as follows. In Section 6.2, we survey hierarchical traffic grooming techniques for networks with a special topology. In Section 6.3, we present a hierarchical framework for traffic grooming in

networks of general topology with either static or dynamic traffic. We present a performance study of hierarchical grooming in Section 6.4, and we conclude the chapter in Section 6.5.

6.2 Hierarchical Grooming in Special Topology Networks

6.2.1 Ring Networks

Early research in traffic grooming focused on ring topologies [2, 3, 4], mainly due to the practical importance of upgrading the existing SONET/SDH infrastructure to support multiple wavelengths. A point-to-point WDM ring is a straightforward extension of a SONET/SDH network, but requires that each node be equipped with one add-drop multiplexer (ADM) per wavelength. Clearly, such a solution has a high ADM cost and cannot scale to more than a few wavelengths. Therefore, much of the research in this context has been on reducing the number of ADMs by grooming sub-wavelength traffic onto lightpaths that optically bypass intermediate nodes, and several near-optimal algorithms have been proposed [3, 4]. However, approaches that do not impose a hierarchical structure on the ring network may produce traffic grooming solutions, in terms of the number of ADMs and their placement, that can be sensitive to the input traffic demands.

The study in [2] was the first to present several hierarchical ring architectures and to characterize their cost in terms of the number of ADMs (equivalently, electronic transceivers or ports) and wavelengths for non-blocking operation under a model of dynamic traffic. In a single-hub ring architecture, each node is directly connected to the hub by a number of lightpaths, and all traffic between non-hub nodes goes through the hub. In a double-hub architecture, there are two hub nodes diametrically opposite to each other in the ring. Each node is connected to both hubs by direct lightpaths, and non-hub nodes send their traffic to the hubs for grooming and forwarding to the actual destination.

A more general hierarchical ring architecture was also proposed in [2]. In this architecture, shown in Fig. 6.1, ring nodes are partitioned into two types: *access* and *backbone*. The set of wavelengths is also partitioned into access and backbone wavelengths. The access wavelengths are used to connect all nodes, including access and backbone nodes, in a point-to-point WDM ring that forms the first level of the hierarchy. At the second level of the hierarchy, the backbone wavelengths are used to form a point-to-point WDM ring among the backbone nodes only. This hierarchy determines the routing of traffic between two access nodes as follows. If the two access nodes are such that there is no backbone node along the shortest path between them, their traffic is routed using single-hop lightpaths over the access ring along the shortest path. Otherwise, suppose that b_1 and b_2 are the first and last backbone nodes, respectively, along the shortest path between two access nodes a_1 and a_2 (note that b_1 and

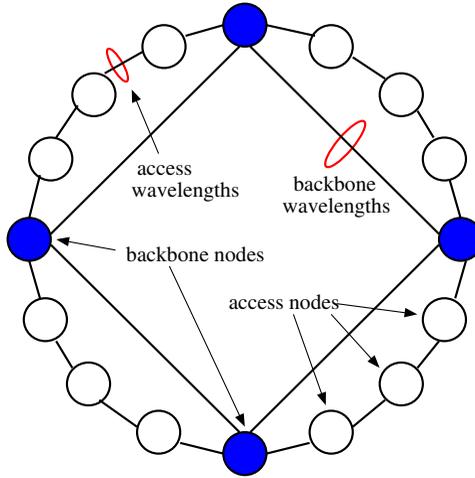


Fig. 6.1. Hierarchical ring architecture with 12 access and 4 backbone nodes

b_2 may coincide). Then, traffic from a_1 to a_2 is routed to b_1 over the access ring, from there to b_2 over the backbone ring, and finally over the access ring to a_2 .

A similar hierarchical ring structure was considered in [11], and it was shown that, using local (access) and bypass (backbone) wavebands, P -port dynamic traffic (in which each node is allowed to send and receive at most P wavelengths worth of traffic) can be supported with a minimum number of wavelengths.

A different hierarchical approach for grooming sub-wavelength traffic in ring networks was introduced in [12]. Specifically, the N ring nodes are grouped into K *super-nodes*, where each super-node consists of several consecutive ring nodes, as shown in Fig. 6.2. The idea behind this partitioning is to pack (groom) all traffic from some super-node x to another super-node y onto lightpaths that are routed directly between the two super-nodes, bypassing intermediate nodes and hence, reducing the number of ADMs required. The study considered both uniform and distance-dependent traffic patterns, and, for each pattern, derived the number K of super-nodes, as a function of the number N of ring nodes and the granularity $C \geq 1$ of each wavelength, so as to minimize the number of ADMs; the granularity C is the number of unit traffic components that can be carried on a single wavelength.

Finally, [2] also proposes the decomposition of a ring into contiguous segments; these are similar to the super-nodes of [12] but are referred to as *subnets*. The ring network is organized in a hierarchical manner as a tree of subnets, where the root of the tree corresponds to a segment that consists of the entire ring. A tree node corresponding to a non-empty subnet s may be subdivided recursively into contiguous subsegments (subnets), and these become the children of subnet s in the tree. The set of wavelengths is also recursively partitioned into *transit* and *internal* sets at each node. Internal

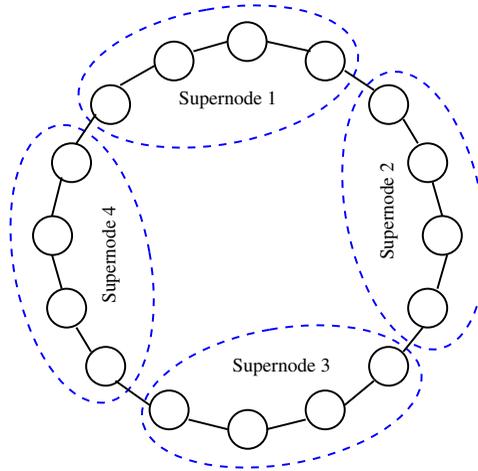


Fig. 6.2. Ring architecture with 4 super-nodes, each of size 4

wavelengths are used within each subnet child of a node to carry traffic local to this subnet, whereas transit wavelengths are used to carry traffic between the subnet children of a node.

6.2.2 Torus, Tree, and Star Networks

A hierarchical approach for networks with a torus or tree topology was presented in [11], and is based on embedding rings on the underlying topology and then selecting hub nodes along each ring and using bypass wavelengths to interconnect the hubs. Consider first a $N \times M$ torus network, whose nodes are logically arranged on a grid of N rows and M columns. The network is viewed as a collection of N row-rings and M column-rings, and several nodes on each ring are designated as hubs; the hub selection is performed using an algorithm described in [11]. Traffic demands from some source s to a destination d are routed in three steps: from s to a hub h_1 in the same row as s along the appropriate row-ring; from h_1 to a hub h_2 in the same column as h_1 along the column-ring; and finally from h_2 to the destination d in the same row along a row-ring. This approach imposes a two-level hierarchy with non-hub nodes at the first level and hub nodes at the second level.

For tree networks, [11] proposes to embed a *virtual ring* in two steps: (1) using depth-first search to visit every node in the tree, and (2) locally arranging the tree nodes in a ring such that the nodes are connected in the ring in the order in which the corresponding tree nodes were first visited by the depth-first search. By defining hubs along the virtual ring, traffic components can be routed using the same algorithm we described for the hierarchical ring in Fig. 6.1.

A traffic grooming algorithm for networks with a star topology was developed in [13]. The algorithm starts by creating lightpaths between the hub and each non-hub node s to carry all traffic originating and terminating at s . Such a solution provides maximum flexibility in terms of grooming, since traffic can be packed efficiently for transmission to the hub, and it can be groomed effectively there for forwarding to the destination. However, it usually requires a large number of lightpaths (equivalently, electronic ports). The algorithm then considers all traffic components in decreasing order of magnitude. Let t be a traffic component from some node s to another node d . The algorithm creates a direct lightpath from s to d to carry t , if there is an available wavelength for doing so; otherwise, no such lightpath is created. A direct lightpath is optically switched at the hub, bypassing electronic switching and grooming, and creating one has the potential to decrease the number of lightpaths by eliminating two lightpaths to/from the hub. The algorithm proceeds until all traffic components have been considered, and returns the solution with the minimum number of lightpaths. It was shown in [13] that this solution is close to optimal for a wide range of problem instances.

6.3 Hierarchical Grooming in General Topology Networks

All the approaches we have discussed so far were developed for networks with topologies that are either symmetric (i.e., ring or torus) or contain no cycles (i.e., tree or star). In this section, we describe a framework for hierarchical traffic grooming that is applicable to networks with a general topology. The framework can be used for static or dynamic traffic, and for either sub-wavelength demands (to be groomed into lightpaths) or full-wavelength demands (to be groomed into wavebands). Although our discussion will consider only two levels of hierarchy, this approach can be extended in a straightforward manner to three or more levels of hierarchy to deal with networks of large size.

The traffic grooming problem involves the following conceptual subproblems (SPs) for sub-wavelength demands [1]:

1. *logical topology SP*: find a set of lightpaths to carry the offered traffic;
2. *traffic routing SP*: route the traffic components over the lightpaths; and
3. *lightpath routing and wavelength assignment (RWA) SP*: assign a wavelength and path over the physical topology to each lightpath.

This is only a conceptual decomposition that helps in understanding and reasoning about the problem; in an optimal approach, the subproblems would be considered jointly in the solution. The first and second subproblems together constitute the grooming aspect of the problem.

The hierarchical grooming approach, first described in [14], emulates the hub-and-spoke model used by the airline industry to “groom” passenger traffic onto connecting flights. Specifically, the network is first partitioned into clusters (or islands) of nodes, where each cluster consists of nodes in a

contiguous region of the network. The clusters form the first level of the hierarchy, and may either correspond to independent administrative entities (e.g., autonomous systems), or may be created solely for the purpose of simplifying resource management and control functions (e.g., as in partitioning a single OSPF administrative domain into multiple areas). Within each cluster, one node is designated as the *hub*, and is responsible for grooming intra-cluster traffic as well as inter-cluster traffic originating or terminating locally. Hub nodes collectively form the second level of the hierarchy, and are expected to be provisioned with more resources (e.g., larger number of switching ports and higher capacity for grooming traffic) than non-hub nodes. Returning to the airline analogy, a hub node is similar in function to airports that serve as major hubs; these airports are typically larger than non-hub airports, in terms of both the number of gates (“ports”) and physical space (for “switching” passengers between gates).

To illustrate this approach, let us consider the 32-node network in Fig. 6.3. The figure shows a partition of the network into eight clusters, B_1, \dots, B_8 , each cluster consisting of four nodes. These clusters represent the first level of the hierarchy. Within each cluster, one node is the hub; for instance, node 2 is the hub for cluster B_1 . The hub nodes of the eight first-level clusters form the second level of the hierarchy, and are responsible for grooming and routing inter-cluster traffic.

The main idea behind the hierarchical grooming strategy is to solve the first and second subproblems of the traffic grooming problem (i.e., construct the logical topology and determine the routing of traffic components on it) separately for each level of the hierarchy. In the first step, each cluster is considered independently of the others, and a set of lightpaths is created to route local (intra-cluster) traffic, as well as inter-cluster traffic to and from

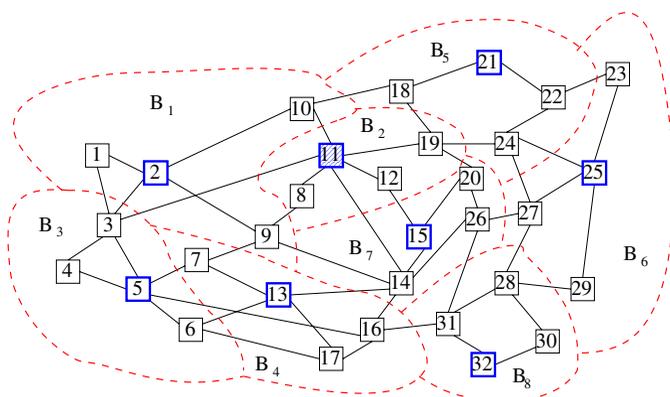


Fig. 6.3. A 32-node network, partitioned into eight first-level clusters B_1, \dots, B_8 , with the corresponding hubs at the second level of the hierarchy

the local hub. In the second step, lightpaths are created between the hub nodes to carry the inter-cluster traffic. Consequently, the problem of routing inter-cluster demands is divided into three simpler subproblems: routing the component to the local hub, from there to the remote hub, and then to the ultimate destination. Finally, given the set of inter- and intra-cluster lightpaths, the third subproblem can be solved on the underlying physical topology of the network using a standard RWA algorithm.

The hierarchical grooming algorithm for sub-wavelength demands consists of three phases:

1. **Clustering and hub selection.** In this phase, the network is partitioned into clusters and one node in each cluster is designated as the hub.
2. **Hierarchical logical topology formation and traffic routing.** During this phase, the first and second subproblems of the traffic grooming problem are solved in an integrated manner. The outcome of this phase is a set of lightpaths for carrying the traffic demands, and a routing of individual traffic components over these lightpaths.
3. **Routing and wavelength assignment.** Each of the lightpaths in the logical topology are assigned a wavelength and path on the underlying physical topology of the network.

This approach has the following desirable characteristics:

- it is hierarchical, facilitating control, management, and security functions;
- it decouples the grooming of traffic components into lightpaths from the routing and wavelength assignment for these lightpaths: grooming is performed on a logical hierarchy of clusters while RWA is performed directly on the underlying physical topology;
- it provisions only a few nodes (the hubs) for grooming traffic that they do not originate or terminate;
- it handles efficiently small traffic demands: at the first level of hierarchy, nodes pack their traffic on lightpaths to the local hub; at the second level, demands among remote clusters are packed onto lightpaths between the corresponding hubs; and
- it allows for large traffic components to be routed on direct lightpaths, eliminating the cost of terminating and switching them at intermediate nodes.

The following subsections discuss each of the three phases of the algorithm in more detail.

6.3.1 Clustering and Hub Selection

The objective of this phase is twofold: (1) to partition the network nodes into some number k of clusters, denoted B_1, \dots, B_k , and (2) to select one node in each cluster B_i as the hub, denoted h_i . Clearly, the number of clusters,

their composition, and the corresponding hubs must be selected in a way that helps achieve the goal of minimizing the number of lightpaths and wavelengths required to carry the traffic demands. Therefore, the selection of clusters and hubs is a complex and difficult task, as it depends on both the physical topology of the network and the traffic matrix T . To illustrate this point, consider the trade-offs involved in determining the number k of clusters.¹ If k is very small (but greater than one), the amount of inter-cluster traffic generated by each cluster will likely be large. Hence, the k hubs may become bottlenecks, resulting in a large number of ports at each hub and possibly a large number of wavelengths (since many lightpaths may have to be carried over the fixed number of links to/from each hub). On the other hand, a large value for k implies a small number of nodes within each cluster. In this case, the amount of intra-cluster traffic will be small, resulting in inefficient grooming (i.e., a large number of lightpaths); similarly, at the second-level cluster, $O(k^2)$ lightpaths will have to be set up to carry small amounts of inter-cluster traffic.

It was observed in [15, 16] that the clustering and hub selection subproblem bears similarities to the classical k -center problem [17, 18]. The objective of the k -center problem is to find a set S of k nodes (centers) in the network, so as to minimize the maximum distance from any network node to the nearest center. Thus, the set S implicitly defines k clusters with corresponding hub nodes in S . A solution to the k -center problem may be useful for hierarchical traffic grooming since it is likely to lead to short lightpaths within a cluster, thus requiring fewer wavelengths. Also, this type of clustering tends to avoid physical topologies with a large diameter for each cluster; such topologies are not a good match for hierarchical grooming that requires nodes to send their traffic to the hub.

The k -center problem is NP-Complete, and the 2-approximation algorithm of [17] was used in [15, 16], with one modification. The modification was based on the observation that the k -center problem takes only the physical topology as input, and its only goal is to minimize the maximum node-to-hub distance; in the traffic grooming context, on the other hand, hub capacity should also be considered. Since hubs are responsible for originating and terminating a larger number of lightpaths than non-hub nodes, it is generally desirable to select as hubs the nodes with the largest bandwidth capacity, i.e., those with the largest physical degree, so as to keep the wavelength requirements for the network low. Therefore, whenever the algorithm of [17] selects a new hub arbitrarily among a set of candidate nodes, it was modified to select the candidate node with the maximum physical degree.

A clustering algorithm designed specifically for hierarchical traffic grooming was presented in [19]. This work identified several grooming-specific factors affecting the selection of clusters and hubs, and developed a parameterized

¹ Note that in the special case of $k = 1$, there is a single cluster with one hub and $N - 1$ non-hub nodes, whereas in the special case $k = N$, there are N clusters, each with a single hub and no non-hub nodes.

algorithm that can achieve a desired trade-off among various goals. The algorithm partitions the network into clusters by considering: (1) the intra- and inter-cluster traffic, attempting to cluster together nodes with “dense” traffic in order to reduce the number of long inter-cluster lightpaths; (2) the capacity of the cut links connecting each cluster to the rest of the network, selecting clusters with a relatively large cut size so as to keep the number of wavelengths low; and (3) the physical shape of each cluster, attempting to avoid clusters with a large diameter. The algorithm also selects hubs on the basis of their physical degree to prevent hub links from becoming bottlenecks. It was shown in [19] that this algorithm outperforms the k -center algorithm in terms of both the port and wavelength costs.

6.3.2 Hierarchical Logical Topology Formation and Traffic Routing

The formation of the hierarchical logical topology for traffic grooming follows three steps: formation of direct lightpaths, intra-cluster lightpaths, and inter-cluster lightpaths. The following discussion assumes that traffic demands are static; at the end of the section, we explain how a hierarchical topology may be formed for dynamic traffic.

Direct lightpaths for large traffic demands. During this step, “direct-to-destination” lightpaths are created between two nodes that exchange large amounts of traffic, even if these nodes belong to different clusters. Similarly, “direct to/from remote hub” lightpaths are created between some node s and a remote hub h if there is a sufficiently large amount of traffic between s and the nodes in h ’s cluster. Setting up such lightpaths to bypass the local and/or remote hub node has several benefits: the number of lightpaths in the logical topology is reduced, the number of ports and switching capacity required at hub nodes is reduced (leading to higher scalability), and the RWA algorithm may require fewer wavelengths (since hubs will be less of a bottleneck).

Intra-cluster lightpaths. At this step, each cluster is considered independently of the others, and intra-cluster lightpaths are formed by viewing each cluster as a *virtual star* (despite the fact that, in general, the actual topology of a cluster is very different than that of a *physical star*, as Fig. 6.3 illustrates). Consider some cluster B with hub h . The intra-cluster lightpaths within cluster B are formed by (1) having all traffic to (respectively, from) any node s of B from (to) nodes outside the cluster originate (terminate) at the hub h , and (2) applying the algorithm for star networks discussed in Section 6.2.2 to cluster B in isolation. Having all inter-cluster traffic originate or terminate at the hub imposes a hierarchical structure to the logical topology of lightpaths: inter-cluster traffic, other than that carried by direct lightpaths set up earlier, is first carried to the local hub, groomed there with inter-cluster traffic from other local nodes, carried on lightpaths to the destination hub (as we discuss shortly), groomed there with other local and non-local traffic, and finally carried to the destination node.

At this stage, the lightpaths to be created are simply identified; the routing of these lightpaths over the physical topology is performed during the RWA phase discussed in the next section. Depending on the actual topology of the cluster B , which may be quite different than that of a physical star, once routed, these lightpaths may follow paths that do not resemble at all the paths of a physical star. Also, recall that the lightpaths created by the star algorithm are either “single-hop” (i.e., from a non-hub node to the hub, or vice versa) or “two-hop” (i.e., from one non-hub node to another). Hence, the routing of the individual traffic components is implicit in the hierarchical logical topology of each cluster.

Inter-cluster lightpaths. At the end of the intra-cluster grooming step, all traffic (other than that carried by the initial direct lightpaths) from the nodes of a cluster B with destination outside the cluster, is carried to its hub h for grooming and transport to the destination hub. In order to carry this traffic, a second-level cluster is considered, consisting of the k hub nodes of the first-level clusters. This cluster is also viewed as a *virtual star* with an associated traffic matrix representing the inter-cluster demands only. The inter-cluster lightpaths to carry these demands are then obtained by applying the star algorithm of Section 6.2.2 to this cluster in isolation. As with intra-cluster lightpaths, the routing of the inter-cluster lightpaths is performed on the underlying physical topology during the RWA phase.

Dynamic Traffic. While the above discussion assumed a static traffic scenario, the three-step approach to forming the logical topology can be adapted to accommodate dynamic traffic. If a new connection request is for a sufficiently large traffic demand, then a direct lightpath is set-up, otherwise grooming must be considered. For an intra-cluster connection request that can be accommodated (groomed) on existing intra-cluster lightpaths (either directly to the destination or through the local hub), no changes in the logical topology are required; otherwise one or two lightpaths will need to be created (from the source to the local hub and/or from the hub to the destination). Similarly, for an inter-cluster request that cannot be accommodated on the current logical topology, up to three new lightpaths may have to be setup (from the source to the local hub, from there to the remote hub, and finally to the destination). We emphasize again that any lightpaths that need to be created are simply identified in this phase; the routing of these new lightpaths is discussed next.

6.3.3 Routing and Wavelength Assignment

For static traffic demands, the outcome of the logical topology phase is a set of lightpaths and an implicit routing of the original traffic components over these lightpaths. In this case, the objective of this phase is to route the lightpaths over the underlying physical topology, and color them using the minimum number of wavelengths. The static RWA problem on arbitrary network topologies has been studied extensively in the literature [8, 9, 20, 21, 22], and any existing algorithm may be used in this case. With

dynamic traffic demands, new lightpaths need to be added to the logical topology each time an arriving connection request cannot be accommodated. An existing dynamic RWA algorithm [23, 24] may be used in this case to find a path and wavelength for the new lightpaths. Hence, by decoupling the grooming and routing of sub-wavelength traffic components onto lightpaths from the routing and wavelength assignment for these lightpaths, hierarchical grooming may capitalize on the the vast body of research on RWA algorithms.

6.3.4 Extension to Lightpath Grooming

The above hierarchical approach may also be applied to networks with multi-granular optical switching capabilities, in which multiple wavelengths may be groomed into wavebands and all wavelengths in a waveband be switched as a group [25]. Let us define a *bandpath*, a generalization of the lightpath concept, as a (waveband, path) pair that uniquely identifies the path over which the set of wavelengths included in the waveband will travel. It was observed in [26] that the traffic grooming problem on a set of full-wavelength demands involves the following three subproblems that are similar to the ones we described earlier:

1. *logical topology SP*: find a set of bandpaths to carry the offered full-wavelength (lightpath) traffic;
2. *lightpath routing SP*: route the lightpaths over the bandpaths; and
3. *bandpath routing and wavelength assignment (RWA) SP*: assign a waveband and path over the physical topology to each bandpath.

Therefore, the hierarchical grooming algorithm can be applied to this problem with only small modifications. Specifically, after partitioning the network and assigning a hub to each cluster, a hierarchical logical topology can be formed by creating bandpaths (instead of single lightpaths) (1) from each node to its local hub; (2) between hubs to carry inter-cluster traffic; and (3) from each hub to the nodes in its cluster. Finally, an existing (static or dynamic) RWA algorithm may be used to assign a waveband and path to each bandpath; since each waveband carries a unique set of wavelengths, assigning a waveband to each bandpath implicitly assigns a wavelength to each lightpath in the bandpath. It was shown in [26] that this hierarchical approach is effective, scalable, and outperforms an existing algorithm for forming and routing bandpaths.

6.4 Performance of Hierarchical Grooming

We now present a small set of experimental results to illustrate the performance of the hierarchical grooming algorithm for static traffic demands. The following methodology was employed in this study. First, the modified k -center algorithm (refer to Section 6.3.1) was used on the 32-node, 53-link

network shown in Fig. 6.3 to obtain three different clusterings with two, four, and eight clusters, respectively; the special case of a single cluster comprising all network nodes was also considered. A random traffic pattern was assumed, and thirty problem instances (i.e., random traffic matrices) were generated. The hierarchical logical topology for each of the 120 instance-clustering pairs was determined by applying the methodology described in Section 6.3.2. The LFAP RWA algorithm [21] was used to route and assign a wavelength to each lightpath; LFAP is fast, conceptually simple, and has been shown to use a number of wavelengths that is close to the lower bound. In order to characterize the performance of the hierarchical grooming solutions, lower bounds on the number of lightpaths and wavelengths necessary to carry a given traffic demand matrix were also obtained; these bounds were obtained independent of the manner (hierarchical or otherwise) in which grooming is performed, as explained in [15].

In order to compare results among different problem instances, two performance metrics were defined: the *normalized lightpath count* and the *normalized wavelength count*. For a given problem instance, the normalized lightpath count is computed as the ratio $lp_h/lp_l \geq 1$, where lp_l is the lower bound on the number of lightpaths, and lp_h is the actual number of lightpaths required in the hierarchical grooming solution. Clearly, the closer this value is to one, the closer the hierarchical solution is to the optimal. The normalized wavelength count is computed in a similar manner.

Figures 6.4 and 6.5 and Table 6.1 present experimental results for the random traffic pattern. Figures 6.4 and 6.5 plot the normalized lightpath and wavelength count, respectively, for each problem instance and corresponding clustering, while Table 6.1 presents aggregate statistics over all 30 problem instances regarding the average lightpath length, the average maximum hub degree (i.e., the maximum of the number of incoming or outgoing wavelengths at the hub), and the average number of wavelengths.

We observe that as the number of clusters into which the network is partitioned increases, the total number of lightpaths in the resulting topology increases gradually (Fig. 6.4). On the other hand, the number of required wavelengths generally decreases as the number of clusters increases (Fig. 6.5), and so do the average lightpath length and the maximum hub degree. These results can be explained by noting that, as the number of clusters increases, the size of each cluster decreases. With a smaller cluster size, more lightpaths are necessary for both intra-cluster traffic (since the amount of traffic within a cluster is relatively small and lightpaths are not utilized efficiently) and inter-cluster traffic (since each hub has to establish lightpaths to a larger number of hubs in other clusters). Also, intra-cluster lightpaths are shorter when clusters are small, and these short lightpaths are less likely to share links, resulting in fewer wavelengths. At the same time, there is relatively less traffic to be groomed at each hub, hence hub degrees (and hub cost) decrease; the fact that hubs are less of a bottleneck also reduces the wavelength requirements.

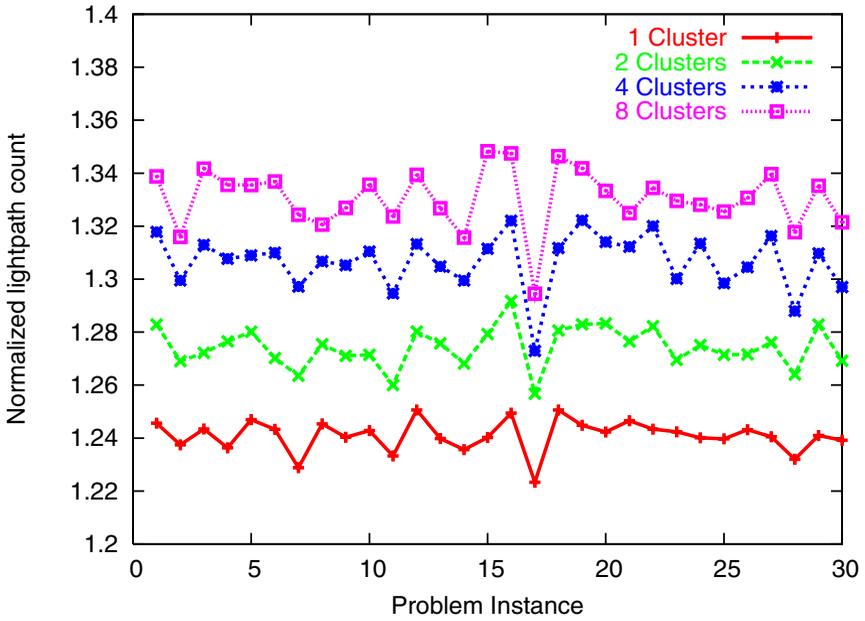


Fig. 6.4. Lightpath comparison

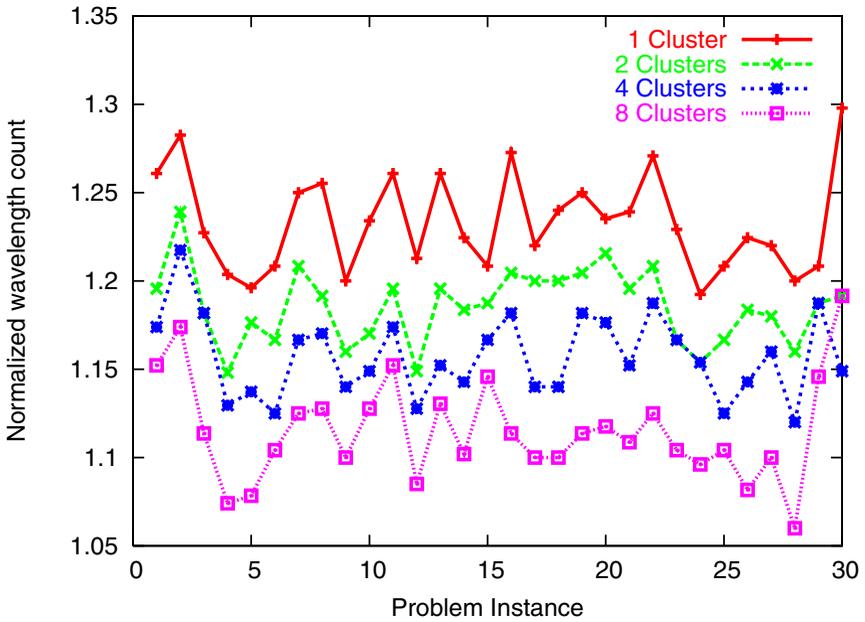


Fig. 6.5. Wavelength comparison

Table 6.1. Aggregate statistics over all 30 instances

#Clusters	Avg Lightpath Length	Avg Max Hub Degree	Avg #Wavelengths
1	3.17	266	60
2	2.93	231	57
4	2.87	182	56
8	2.75	145	53

From Fig. 6.4, we note that the number of lightpaths created by the hierarchical grooming approach are only about 25–35% above the lower bound, and this behavior is consistent across all problem instances. From Fig. 6.5, we observe that, with appropriate clustering, the wavelength requirements of this approach are close to the lower bound. Recall that both lower bounds have been computed in a manner that is independent of the grooming methodology employed. Consequently, these results demonstrate that, despite its hierarchical nature, this approach produces grooming solutions that are close to optimality.

Similar results regarding the performance of the hierarchical grooming algorithm can be found in [15, 19, 26]. Overall, these results demonstrate that hierarchical traffic grooming can be applied efficiently to large size networks, and produces logical topologies whose lightpath and wavelength requirements are close to the corresponding lower bounds.

6.5 Conclusion

Hierarchical traffic grooming is an efficient and scalable approach to grooming multigranular traffic in large-scale WDM networks with a general topology. We presented a hierarchical framework that is applicable for both static and dynamic contexts, and for either sub-wavelength or full-wavelength traffic. The framework generalizes earlier techniques developed for networks with a special topology (e.g., ring, torus, or tree), and consists of three phases: clustering and hub selection, hierarchical logical topology design, and routing and wavelength assignment. Hierarchical grooming has been shown to perform well over a range of network topologies and traffic patterns, and scales to networks of realistic size.

References

1. R. Dutta and G. N. Rouskas. Traffic grooming in WDM networks: Past and future. *IEEE Network*, 16(6):46–56, November/December 2002.
2. O. Gerstel, R. Ramaswami, and G. Sasaki. Cost-effective traffic grooming in WDM rings. *IEEE/ACM Transactions on Networking*, 8(5):618–630, October 2000.

3. P.-J. Wan, G. Calinescu, L. Liu, and O. Frieder. Grooming of arbitrary traffic in SONET/WDM BLSRs. *IEEE Journal on Selected Areas in Communications*, 18(10):1995–2003, 2000.
4. R. Dutta and G. N. Rouskas. On optimal traffic grooming in WDM rings. *IEEE Journal on Selected Areas in Communications*, 20(1):110–121, January 2002.
5. V. R. Konda and T. Y. Chow. Algorithm for traffic grooming in optical networks to minimize the number of transceivers. In *IEEE Workshop on High Performance Switching and Routing*, pp. 218–221, 2001.
6. D. Li, Z. Sun, X. Jia, and K. Makki. Traffic grooming on general topology WDM networks. *IEE Proceedings on Communications*, 150(3):197–201, June 2003.
7. K. Zhu and B. Mukherjee. Traffic grooming in an optical WDM mesh network. *IEEE Journal on Selected Areas in Communications*, 20(1):122–133, Jan 2002.
8. J.Q. Hu and B. Leida. Traffic grooming, routing, and wavelength assignment in optical WDM mesh networks. *Proceedings of the IEEE INFOCOM 2004*, 23(1):495–501, March 2004.
9. Z.K.G. Patrocínio Jr. and G. R. Mateus. A lagrangian-based heuristic for traffic grooming in WDM optical networks. In *IEEE GLOBECOM 2003*, 5(1–5): 2767–2771, Dec 2003.
10. C. Lee and E. K. Park. A genetic algorithm for traffic grooming in all-optical mesh networks. In *IEEE SMC 2002*, 7, 6–9, Oct 2002.
11. L.-W. Chen and E. Modiano. Dynamic routing and wavelength assignment with optical bypass using ring embeddings. *Optical Switching and Networking*, 1(1):35–49, 2005.
12. J. Simmons, E. Goldstein, and A. Saleh. Quantifying the benefit of wavelength add-drop in WDM rings with distance-independent and dependent traffic. *IEEE/OSA Journal of Lightwave Technology*, 17:48–57, Jan 1999.
13. B. Chen, R. Dutta, and G. N. Rouskas. Traffic grooming in star networks. In *Proceedings of the First Traffic Grooming Workshop*, October 2004.
14. B. Chen, G. N. Rouskas, and R. Dutta. A framework for hierarchical grooming in WDM networks of general topology. In *Proceeding of BROADNETS 2005*, pp. 167–176, October 2005.
15. B. Chen, G. N. Rouskas, and R. Dutta. On hierarchical traffic grooming in WDM networks. *IEEE/ACM Transactions on Networking*, 16(6), December 2008.
16. B. Chen, R. Dutta, and G. N. Rouskas. On the application of k -center algorithms to hierarchical traffic grooming. In *Proceedings of the Second Traffic Grooming Workshop*, pp. 295–301, October 2005.
17. T. Gonzalez. Clustering to minimize the maximum inter-cluster distance. *Theoretical Computer Science*, 38:293–306, 1985.
18. D. Hochbaum and D.B. Shmoys. A best possible heuristic for the k -center problem. *Mathematics of Operations Research*, 10:180–184, 1985.
19. B. Chen, G. N. Rouskas, and R. Dutta. Clustering and hierarchical traffic grooming in large scale WDM networks. In *Proceedings of ONDM 2007*, May 2007.
20. I. Chlamtac, A. Ganz, and G. Karmi. Lightpath communications: An approach to high bandwidth optical WANS. *IEEE Transactions on Communications*, 40(7):1171–1182, July 1992.
21. H. Siregar, H. Takagi, and Y. Zhang. Efficient routing and wavelength assignment in wavelength-routed optical networks. *Proc. 7th Asia-Pacific Network Oper. and Mgmt Symposium*, pp. 116–127, Oct. 2003.

22. S. Baroni and P. Bayvel. Wavelength requirements in arbitrary connected wavelength-routed optical networks. *IEEE/OSA Journal of Lightwave Technology*, 15(2):242–251, Feb. 1997.
23. X-W. Chu, B. Li, and Z. Zhang. A dynamic RWA algorithm in a wavelength-routed all-optical network with wavelength converters. In *Proceedings of IEEE INFOCOM*, pp. 1795–1802, 2003.
24. A. Mokhtar and M. Azizoglu. Adaptive wavelength routing in all-optical networks. *IEEE/ACM Transactions on Networking*, 6(2):197–206, April 1998.
25. X. Cao, V. Anand, and C. Qiao. Waveband switching in optical networks. *IEEE Communications Magazine*, 41(4):105–112, 2003.
26. M. Iyer, G. N. Rouskas, and R. Dutta. A hierarchical model for multigranular optical networks. In *Proceedings of IEEE BROADNETS 2008*, September 2008.

Part II

Techniques

Traffic Grooming in SONET/SDH Rings

Raza Ul-Mustafa

7.1 Introduction

This chapter deals with the problem of traffic grooming on WDM ring networks employing SONET equipment. The chapter addresses the problems of dimensioning and provisioning such networks in order to support static traffic. We briefly review the pertinent features of SONET rings in the next paragraph; more details are available in Chapter 4 and the references cited there in.

SONET rings, which are very widely used in Metro networks, are of two types:

- Unidirectional Path-Switched Ring (UPSR), Fig. 7.1(a), and
- Bidirectional Line-Switched Ring (BLSR), Fig. 7.1(b).

UPSR rings use two fibers for transmission in two opposite directions, where one fiber is the working fiber and the other is the protection fiber. The signal from a source is transmitted on both fibers, and the receiver monitors both of them, and then selects the better of the two signals. BLSR rings use either two or four fibers, hence called BLSR/2 and BLSR/4, respectively. In the BLSR/4 ring (shown in Fig. 7.1(b)), two fibers are used as working fibers, and the other two are used for protection. However, the working traffic can be carried in either of the two directions (typically using the shortest path), and the protection traffic is transmitted in the opposite direction, only in the case of failure. BLSR/2 is similar in topology to UPSR, but is similar in operation to BLSR/4. The use of two fibers in BLSR/2 requires that the capacity of each fiber be split between working and protection bandwidth.

SONET rings in metro networks are typically provisioned to support static traffic. Under static traffic, a fixed matrix of traffic demands is available, i.e., all the demands are known a priori. Given a physical topology, the task is to design and provision the network such that the network cost is minimized. The traffic matrix itself can be further classified into one of the following types:

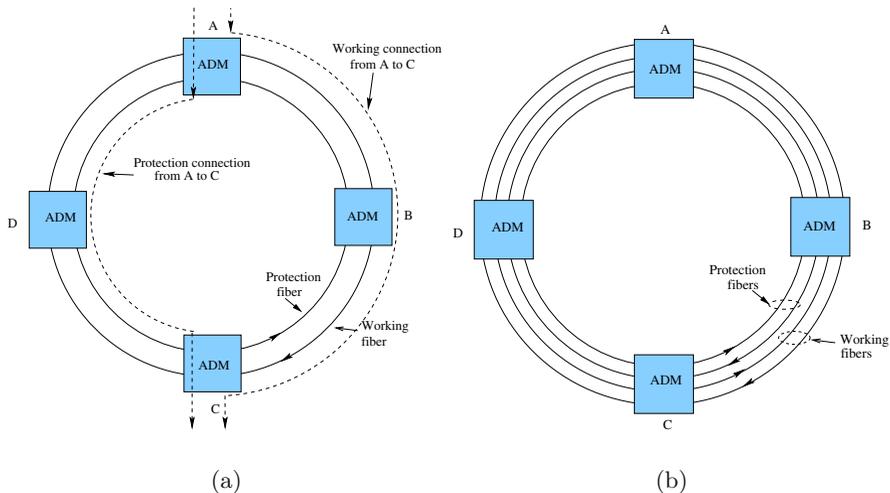


Fig. 7.1. SONET rings: (a) UPSR; (b) BLSR/4

1. *Uniform* traffic demands, where each pair of nodes in the network exchanges the same level of traffic units,
2. *Distance-dependent* traffic demands, where the number of traffic units between a node pair depends on the distance between them (i.e., a node shares more traffic with the neighboring nodes than the nodes farther from it), and
3. *Non-uniform* traffic demands, which is the more general case, where the number of traffic units between node pairs is arbitrary.

We will refer to the two special cases of uniform and distance-dependent traffic patterns, as the *fixed traffic pattern*, and to the case of non-uniform traffic demands as *arbitrary traffic pattern*.

Some of the earliest studies in the area of traffic grooming on ring architecture with static traffic patterns were conducted by Simmons, Goldstein, and Saleh [16, 17, 18], Gerstel, Lin and Sasaki [9, 10], Chiu and Modiano [4, 5], and Zhang and Qiao [24, 25]. All such studies advocated that, to reduce the network cost, more emphasis should be placed on minimizing the number of higher layer components than the number of wavelength channels. However, this is a hard problem and was proven to be NP-Complete by Chiu and Modiano [5] and Li et al. [13] by reduction from the bin packing problem (see Chapter 5 for a detailed discussion on the complexity of the traffic grooming problem). Researchers addressed the problem by providing bounds, heuristics, and exact solutions. Table 7.1 provides a classification of some of the related work in this area.

Table 7.1. Related work in the area of traffic grooming on ring architecture with static traffic patterns

Traffic Pattern	Solution Approach	References
Fixed Traffic	Bounds	[1, 3, 4, 5, 10, 16, 17, 24]
	Heuristics	[3, 4, 5, 16, 17, 24, 25]
	ILPs	[11]
Arbitrary Traffic	Bounds	[8, 13, 21, 22, 23, 25]
	Heuristics	[8, 13, 19, 21, 22, 23, 25, 26]
	ILPs	[8, 11, 22]

7.2 Fixed Traffic Patterns

In this section, we will review the literature related to the design of WDM networks for traffic grooming on ring topologies using fixed traffic patterns.

7.2.1 Bounds for General Ring Topologies

Simmons, Goldstein, and Saleh provided some initial work in [16] and later refined it in [17]. In both references, they considered a ring topology with fixed traffic pattern. Specifically, in [17], the authors considered uniform all-to-all traffic and distance-dependent traffic between nodes connected as a ring topology. Uniform all-to-all traffic is defined as a traffic scenario whereby a single unit of traffic is sent from each node on a ring to every other node on that ring, while distance-dependent traffic is defined as a traffic scenario whereby the number of traffic units sent between nodes increases as the inter-nodal distance decreases. The authors provided upper bounds on the number of SONET Add/Drop Multiplexers (ADMs) and also provided procedures to obtain these bounds. If r represents the number of units of traffic between a source-destination pair and G represents the total number of traffic units that a single wavelength can accommodate,¹ then Equation (7.1) expresses the estimate given in [17] of the required number of ADMs per wavelength for a scenario where each of the N nodes² in a ring is transmitting $1/G$ units of traffic to every other node in that ring. Similarly, Equation (7.2) provides an estimate for the total number of required ADMs for the scenario where each node transmits a full wavelength to each other node, i.e., $r = G$.

$$r = \frac{1}{G} : \text{Number of ADMs per wavelength} = \frac{4N}{N+1} \quad (7.1)$$

$$r = G : \text{Number of ADMs per wavelength} = \frac{N(N-1)}{2} \quad (7.2)$$

¹ The terms *grooming factor* or *grooming ratio* are used interchangeably to refer to G in the rest of this chapter.

² N was assumed to be *odd*.

They also introduced the interesting concept of a *super node*, where nodes are grouped into distinct super nodes such that the amount of traffic between super nodes occupies a full wavelength. The traffic between super nodes can then be routed using procedures developed for full wavelength traffic. The nodes, however, were grouped into super nodes using approximations, and hence the overall solution is an approximate one. In [17] the authors indicated that, to reduce the number of ADMs, more connections need to be routed optically at each node and have therefore quantified the benefit of using wavelength add-drops by measuring the through-to-total ratio of connections at each node in the network. In order to analyze the distance-dependent traffic scenario, they considered a traffic pattern on a bidirectional ring with an odd number N of nodes for which the amount of traffic between the most distant nodes is one unit, and the traffic demand increases by one unit as the inter-nodal distance decreases by one link. They estimated the number of units of traffic sent by each node, in each direction to be $(N^2 - 1)/8$, and the number of traffic units on each link to be $(N^2 - 1)(N + 3)/48$. The minimum number of required ADMs, for $G = 1$, to support this traffic was then estimated as:

$$G = 1 : \text{Minimum number of ADMs} = \frac{N(N^2 - 1)}{8} \quad (7.3)$$

Finally, another important contribution of this work was the realization that minimizing the number of ADMs will not only reduce the cost of the network, but also in some cases, such as IP over WDM, the ADM savings may imply a substantial simplification in packet routing.

7.2.2 Bounds for UPSR and BLSR/2 Rings

In [4], Chiu and Modiano considered all-to-all uniform traffic on WDM unidirectional ring networks. In [5], they extended their work by considering distance dependent-traffic also. One major contribution of this work is the introduction of upper and lower bounds for the required number of wavelengths and ADMs. Furthermore, they proposed heuristic solutions with a performance that is close to the derived bounds. Using the same definitions for r , G and N above, the authors in [5] give the following loose lower bound on the number of required ADMs:

$$\text{Number of ADMs} \geq \left\lceil \frac{(N - 1)r}{G} \right\rceil N \quad (7.4)$$

They also provided two upper bounds on the minimum number of required ADMs, which are shown in Equations (7.5) and (7.6):

$$\text{Minimum number of ADMs} \leq \left\lceil \frac{N(N - 1)r}{2G} \right\rceil N \quad (7.5)$$

$$\text{Minimum number of ADMs} \leq \left\lceil \frac{r}{G} \right\rceil N(N - 1) \quad (7.6)$$

They also proposed two heuristic approaches to accommodate all-to-all uniform traffic while minimizing the required number of ADMs. The first heuristic attempts to maximize the number of nodes that require fewer ADMs. For example, since a node uses k ADMs if it transmits/receives on k wavelengths, then if M_k represents the number of nodes with k ADMs, the heuristic first tries to maximize M_1 , then M_2 , then M_3 , and so on. The second heuristic assigns nodes to wavelengths by attempting to pack the wavelengths efficiently. All of the N nodes are first divided into $N/\lfloor\sqrt{G}\rfloor$ groups. The cross traffic of each pair in a group is then assigned to a wavelength. This is possible because, by design, the cross traffic between two groups of size $\lfloor\sqrt{G}\rfloor$ is less than G . The traffic within each group is first accommodated on existing wavelengths, and then on additional wavelengths if needed. Besides all-to-all-uniform traffic, the authors also discussed the distance-dependent traffic and derived the upper and lower bounds on the number of ADMs shown in Equations (7.7) and (7.8) below, respectively.

$$\text{Number of ADMs} \leq N \left\lceil \frac{N(N^2 - 1)}{8G} \right\rceil \quad (7.7)$$

$$\text{Number of ADMs} \geq \left\lceil \frac{N(N^2 - 1)}{4G} \right\rceil \quad (7.8)$$

Finally, one of the results mentioned in [5] suggests that the number of wavelengths and ADMs cannot always be minimized simultaneously. However, the authors conjectured that for the special case where each source-destination pair has a single unit of traffic between them, the minimum number of ADMs could be achieved with the minimum number of wavelengths. Later, in [1] Bermond and Coudert proved this conjecture to be true for few values of G (e.g., $G = 3$), while proving it to be false for many other values of G (e.g., $G = 7$).

Reference [10] focused on obtaining bounds on the number of SONET ADMs when nodes are placed on either a UPSR or a BLSR. The authors, however, restricted their analysis to uniform traffic between the nodes. They considered single-hub UPSRs, and assumed that low speed traffic can be cross-connected at nodes. Equation (7.9) shows the lower bound on the number of required ADMs when a UPSR ring is considered (with uniform traffic), while Equation (7.10) gives the lower bound when a single-hub UPSR ring is considered.

$$\text{UPSR: Number of ADMs} \geq \max \left\{ \left\lceil \frac{2N(N-1)r}{G+r} \right\rceil, N \right\} \quad (7.9)$$

$$\text{UPSR, hub: Number of ADMs} \geq (N-1) \left\lceil \frac{(N-1)r}{G} \right\rceil + (N-1)k + \lambda_f \quad (7.10)$$

In Equation (8.10), $k = \lfloor (N - 1)r/G \rfloor$ is the number of wavelengths dedicated to each non-hub node, and λ_f is the number of wavelengths used to carry any leftover traffic for non-hub nodes that was not accommodated by the k wavelengths. After comparing their computed lower bound for single-hub UPSR with the one given in [5], which assumes that traffic streams cannot be cross-connected, the authors concluded that using cross-connection usually results in using fewer ADMs. However, the drawback of using a single-hub architecture is the increase in the number of wavelength channels. The authors also demonstrated that although spatial reuse of the bandwidth of BLSR/2 ring networks makes it less expensive to implement than UPSR, in a few cases (for which $r \leq G/2$) UPSR cannot be more expensive than BLSR/2 in terms of the number of ADMs needed to accommodate the traffic. Finally, the authors considered a network architecture with two different line speeds, namely, OC-12 ($G = 4$) and OC-48 ($G = 16$). They plotted the lower bounds on the cost using Equation (7.9), while using a cost-factor of 2.5 for OC-48 over OC-12. The conclusions were that the cost of using OC-48 is often less than the cost of using OC-12.

In reference [3], the authors computed lower bounds on the number of wavelengths in a BLSR network under static all-to-all uniform traffic. They also computed lower bounds on the number of ADMs in unswitched UPSR rings for uniform all-to-all traffic, which are given by Equation (7.11):

$$\begin{aligned} \text{Number of ADMs} &\geq \max \left\{ \left\lceil \frac{N(N-1)(n+f)}{n(n-1)(1-f) + 2\lfloor G/r \rfloor f} \right\rceil, N \right\} \\ n &= \left\lceil r + \frac{\sqrt{r^2 + 8rG}}{2r} \right\rceil \\ f &= \begin{cases} 1, & \text{if } \frac{2\lfloor G/r \rfloor}{n-1} - n \geq 1 \\ 0, & \text{otherwise} \end{cases} \end{aligned} \quad (7.11)$$

For the UPSR case, the authors also presented a heuristic approach that first constructs full circles, and later grooms these circles onto wavelengths. They focused on balancing the ADM utilization at every node and not splitting a traffic stream into more than one wavelength. The circles are constructed in a manner similar to that of [25]. To assign the circles onto wavelengths a multi-phase approach is used. In phase 1, nodes are divided into clusters such that the traffic between all the nodes of a cluster could be accommodated by one wavelength without splitting any stream. After that, the traffic between the nodes from two different clusters is groomed, provided it results in an effective use of ADMs. In phase 2, more circles are groomed onto existing wavelengths provided that a wavelength has enough capacity and the addition of an ADM does not decrease the ‘‘circles per ADM’’ factor (the factor captures the efficiency of ADM utilization). In phase 3, any remaining circles are groomed by adding such circles to a wavelength that share one or more end points.

7.2.3 Algorithms Based on Graph Theory

A number of studies have used graph theoretic approaches to study the problem of SONET ring network provisioning under the case of static traffic. The authors in reference [1] used tools from graph theory and design theory to address the traffic grooming problem in unidirectional WDM ring networks. They also considered uniform all-to-all traffic demands among the nodes of the network. They showed that the problem of minimizing the number of ADMs for uniform all-to-all traffic demands can be expressed by partitioning the edges of a complete graph of N vertices into W subgraphs, where W is the number of wavelengths such that the total number of vertices in each subgraph is minimized. Each subgraph in this case represents a wavelength, while the vertices in each subgraph correspond to ADMs. The authors pointed out that this problem is similar to design theory, such that an $(N, k, 1)$ -design is nothing but a partition of the edges of a graph into subgraphs isomorphic to blocks in this theory. The classical equivalent definition is as follows: *given a set of N elements, find a set of blocks such that each block contains k elements and each pair of elements appears in exactly one block.* Thereafter, using techniques developed in design theory, the authors computed the number of ADMs required for many possible combinations of the grooming factor and the number of nodes in the network.

7.2.4 Algorithms Based on Meta-Heuristics

Some researchers have used meta-heuristics to solve the traffic grooming problem on WDM ring networks. In reference [6], the authors used Simulated Annealing as one part of the solution approach. They considered WDM ring networks with uniform traffic and presented two different approaches. In one approach, the traffic needs to be routed through a hub node in the network. They defined this approach as a *multihop* approach, because each traffic demand delivery needed to employ more than one lightpath hop. Also, they consider the case in which traffic can be routed without going through any hub node and called it *single-hop* approach. For the multihop approach, they first place ADMs on the nodes for each request and then try to re-assign the traffic such that the maximum number of ADMs and wavelengths can be reduced. For the single-hop approach, they employed a two-phase methodology. In the first phase, they constructed the virtual circles, while in the second phase they used Simulated Annealing to group the circles into different wavelengths. Based on simulation results, they argued that it was beneficial in terms of the number of ADMs to use the single-hop approach while employing simulated annealing for small grooming ratios. However, for large grooming ratios and large networks, the multihop approach could lead to better savings in terms of the number of ADMs.

7.2.5 Algorithms for Multi-Ring Topologies

The authors in reference [15] explored an interesting variant of the ring topology. They considered a topology in which many rings are connected to each other forming a WDM multi-ring network. They classified each traffic connection as either *intra-connection* or *inter-connection*. For an intra-connection, the source and destination are located on the same ring, while for an inter-connection the source and destination are located on different rings. Four virtual topologies are formed depending on how inter-connection traffic is transmitted and also how those connections are grouped onto a wavelength. The four topologies are referred to as: independent virtual topology (IVT), separated virtual topology (SVT), mixed virtual topology (MVT), and partial mixed virtual topology (PMVT).

- In IVT, an inter-connection is cut at the inter-ring connection node, and therefore all traffic is considered as intra-connection. Accordingly, at an inter-ring connection node every wavelength needs to have an ADM.
- In SVT, a connection is strictly classified as intra-connection or inter-connection, and a wavelength carries only one type of connection. An inter-ring connection node thus needs ADMs only for the inter-connection wavelengths.
- In MVT, a wavelength could carry inter-connection traffic, intra-connection traffic, or both.
- In PMVT, some intra-connections and inter-connections are groomed together onto a wavelength to form an MVT, while other connections are groomed separately as in SVT.

Finally, to groom the traffic for each type of virtual topology, the authors proposed four heuristics, the independent traffic grooming (ITG), the separated traffic grooming (STG), the mixed traffic grooming (MTG), and the partially mixed traffic grooming (PMTG), for IVT, SVT, MVT and PMVT, respectively. Using simulation, they showed that ITG requires the minimum number of wavelengths but uses the most number of ADMs at an inter-ring connection node, MTG requires the maximum number of wavelengths, while STG and PMTG require the minimum number of ADMs.

7.3 Arbitrary Traffic Patterns

In this section, we will review the literature related to the design of WDM networks for traffic grooming on ring topologies in order to support arbitrary traffic between the different node pairs.

7.3.1 Single-Hub SONET Rings

Li et al., in reference [13] extended the work done in reference [10] by focusing only on single-hub SONET/WDM rings, while considering uniform and non-

uniform traffic. First, they proved that, given a set of traffic demands, the single-hub BLSR/2 costs no more than the single-hub UPSR under any traffic pattern³. They then showed that optimal traffic grooming can be confined to a narrow subset of valid grooming strategies, and referred to these as *canonical groomings*. For uniform traffic, they computed the required number of ADMs in the working fiber of a UPSR ring, $F(G, r, N)$, which is given by Equation (7.12). The total number of required ADMs for a UPSR ring is then $2F(G, r, N)$, and for a BLSR is $F(G/2, r, N)$.

$$F(G, r, N) = N \left\lceil \frac{r}{G} \right\rceil + N \left\lfloor \frac{r}{G} \right\rfloor + \left\lceil \frac{N}{\lfloor G/(r \bmod G) \rfloor} \right\rceil \quad (7.12)$$

For non-uniform traffic, the authors showed that grooming on a single-hub BLSR/2 could be solved using bin-packing approaches. They proved that using a First-Fit-Decreasing (FFD) algorithm [7] to assign non-uniform integer traffic units to ADMs produces a 10/9 approximation ratio. A typical first-fit-decreasing algorithm first sorts the input objects in decreasing order, and then assigns the objects (sequentially) to a first-fit bin. Therefore, the authors first sorted the residual traffic demands, $r_i \bmod G$, at all non-hub nodes decreasingly and then assigned the demands to the first ADM with sufficient capacity. Note that the remaining traffic demands at each node occupy $\lfloor r_i/G \rfloor$ full wavelengths, and each wavelength in the case of the single-hub architecture requires two ADMs. Finally, for non-uniform integer traffic demands and grooming factors of 2, 4, and 8, the authors also presented optimal algorithms by solving an integer bin packing problem. It turns out that in a few cases the solution strategy is similar (but not exactly the same) to that of FFD.

7.3.2 SONET BLSR

In reference [21], the authors considered SONET/WDM BLSRs, while taking into account non-uniform traffic. They assumed each traffic stream to consist of unitary traffic demand. However, each traffic stream can occur between any pair of arbitrary nodes (as opposed to all-to-all or one-to-all traffic). Therefore, such a setting permits non-uniform traffic. They extended the work presented in reference [20] by considering two versions of the minimum ADM cost problem. In the first version, named *arc-version*, each traffic stream has its predetermined routing, such as shortest path routing. Therefore, the two (working) fiber rings⁴ are handled separately, such that in each fiber ring, a traffic stream is represented as a (directed) circular arc. For the second

³ Compare this result with [10] in which they showed that in a few cases UPSR does a little better than (non-hub) BLSR/2. These results do not contradict each other, as the conclusions in [13] are for single-hub rings only, while those of [10] are for non-hub rings.

⁴ In the case of a BLSR/2 ring, each fiber can be regarded as two virtual fibers, one used for working traffic and the other used for protection.

version, named *chord-version*, they assumed that each traffic stream is full-duplex with symmetric and unitary demands, which must be routed along the same path but in the opposite direction. Thus, the two (working) fiber rings are treated as one (undirected), and each traffic stream is treated as a (undirected) chord. In terms of solution approaches, they presented some lower bounds, and also presented a set of heuristics. To express these bounds, let A represent the set of input arcs, and let $\delta_A(i)$ and $\tau_A(i)$ represent the total number of arcs in A that originate (terminate) from (to) node i , respectively. A lower bound on total number of required ADMs is then given by:

$$\text{Number of ADMs} \geq \sum_{i=1}^N \left[\max \left(\delta_A(i), \frac{\tau_A(i)}{G} \right) \right] \quad (7.13)$$

To present a lower bound on the number of ADMs for the chord-version, let B represent the input set of chords and $\theta_B(i)$ represent the total number of chords in B that contain node i as one endpoint. The lower bound on ADMs is then given by:

$$\text{Number of ADMs} \geq \sum_{i=1}^N \left[\frac{\theta_B(i)}{2G} \right] \quad (7.14)$$

For the heuristic approach, the basic idea is to first construct *primitive rings* from the unitary traffic demands and then partition them into a number of groups such that each group consists of a set of primitive rings and the size of each set does not exceed the grooming factor G . Given a set of primitive rings, they propose the following algorithm, named iterative matching, for constructing groups. Let m represent the total number of primitive rings, $m \bmod G = 0$ (which could be achieved by adding dummy rings) and G is a power of 2. Let Π_0 be the original sets. The i th iteration starts with Π_{i-1} , a 2^{i-1} -grouping of Π_0 , and finds a maximum-weighted perfect matching of Π_i . Then, for each edge in the obtained matching, the two sets incident to the edge are merged. Thus, the i th iteration outputs a 2^i -grouping of Π_0 , denoted by Π_i . The authors also extended the iterative matching algorithm for the case when G is not a power of 2. This work, however, did not include any performance study of the heuristics.

7.3.3 Two-Phase Heuristic Approaches

In [25], the authors extended their earlier work in [24] and proposed heuristic approaches to solve the traffic grooming problem on unidirectional and bidirectional rings, while considering both uniform and non-uniform traffic. The main idea of the heuristics is to follow a two-phase approach. In the first phase, *circles* are constructed using traffic streams, and in the second phase those circles are groomed onto the set of the available wavelengths. Three heuristics were presented to construct circles for uniform traffic, and one heuristic for the

construction of circles for non-uniform traffic was introduced. In addition, a heuristic for circle grooming was provided. The heuristic for the construction of circles using non-uniform traffic may be summarized as follows.

1. Construct circles with the traffic streams having the longest stride.
2. For each of the remaining traffic streams: accommodate it on an already constructed circle if it does not overlap with any other traffic stream present on that circle *and* it also shares at least one end node with any of the present traffic streams on that circle; otherwise, add this traffic stream to a GapMaker list.
3. For each traffic stream in the GapMaker list: try to accommodate the traffic stream on already constructed circles if the restrictions mentioned in Step 2 are met; otherwise, create a new circle with this traffic stream.

The heuristic presented for grooming of circles onto wavelengths first determines the number of circles, m_w , to be groomed on wavelength w . After that, for each wavelength w the following two steps are executed:

1. Find the circle which has the maximum number of ADMs over all existing circles and groom it onto wavelength w .
2. Repeat $(m_w - 1)$ times: groom such a circle onto wavelength w , which when groomed onto w , results in a minimum number of additional ADMs.

Another two-phase approach that addresses the traffic grooming problem on unidirectional and bidirectional rings is presented in [19]. For unidirectional rings, the authors first mapped the ring onto an extended linear topology. They defined a collection of traffic streams that do not share any physical link with each other as a *string*. Then, in the first step of their two-phase approach, they allocated the input traffic onto the extended linear topology while minimizing the number of the strings. In the second step, they employed a grouping algorithm to combine the strings into wavelengths while minimizing the total number of ADMs. In Fig. 7.2, a simple five-node unidirectional ring and its mapping onto a linear topology is depicted. Nodes 6, 7, 8, and 9 (also represented as 1', 2', 3', and 4', respectively) are the added dummy nodes, which correspond to nodes 1, 2, 3, and 4, respectively. Traffic sourced at a node and destined to a lower indexed node will now terminate at the corresponding added dummy node. Once all the traffic is mapped from a unidirectional ring onto a linear topology, the solutions developed for the linear topology below will be applicable to the unidirectional ring too. Figure 7.3 shows the output of the algorithm that minimizes the number of strings, when a sample input and a linear topology of 5 nodes is used. The traffic streams are shown in Fig. 7.3(a). Each segment corresponds to a single traffic unit. Figure 7.3(b) shows the traffic streams after sorting. Figure 7.3(c) shows the output of the string minimization algorithm after efficiently combining the non-overlapping traffic streams (as strings). For each string, the circles at nodes represent the streams with common end-points alluding that an ADM could be shared between these streams. The key idea of the grouping algorithm is to group together, into a

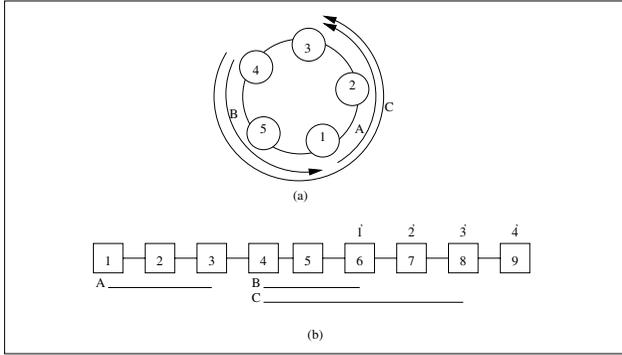


Fig. 7.2. Mapping a unidirectional ring into a linear topology; nodes 6 to 9 are the added dummy nodes corresponding to nodes 1 to 4, respectively

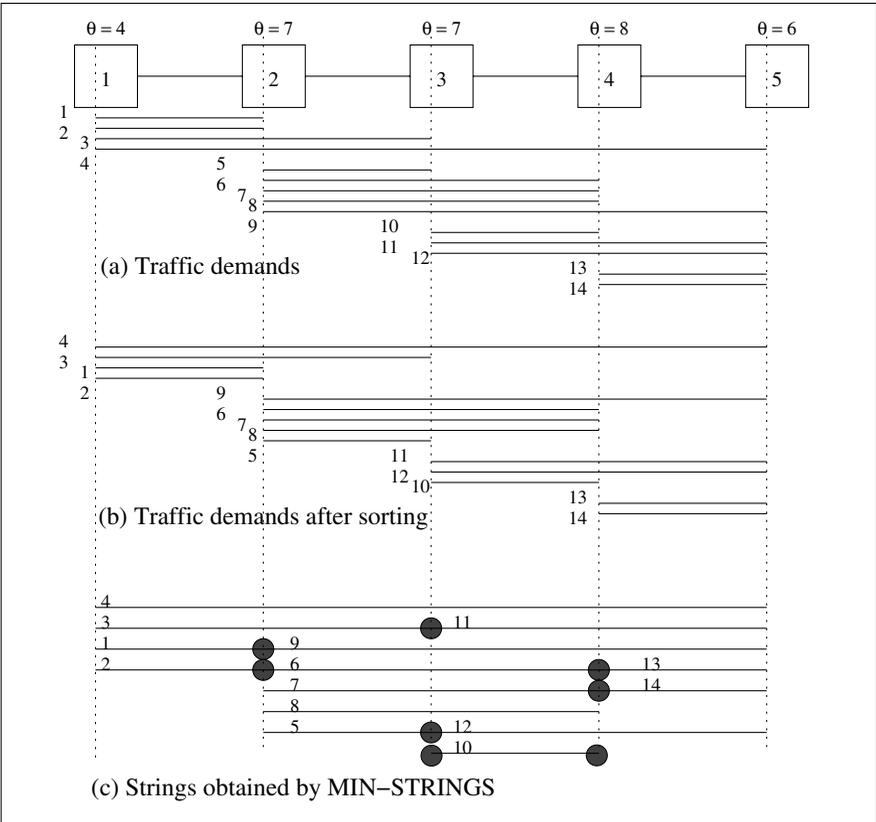


Fig. 7.3. Packing the traffic demands while minimizing the number of strings

wavelength, such strings that share maximum number of common end-points. This leads to sharing of ADMs, resulting in fewer ADMs for that wavelength.

For the bidirectional ring architecture, the authors developed techniques that map a bidirectional ring onto a set of unidirectional rings. They also proposed some traffic routing techniques that reduce the number of ADMs needed to accommodate the input traffic matrix. The time complexity of their techniques is shown to be at least an order of N lower than the other proposed two-phase approaches.

7.3.4 Meta-Heuristic Approaches

Meta-heuristics were also used to solve the SONET ring provisioning problem under arbitrary traffic patterns. For example, in reference [21] the authors presented a Genetic Algorithm-based approach for unidirectional SONET/WDM rings with arbitrary and asymmetric traffic. They chose an interesting representation of the *chromosome* by making use of an order-based approach, which essentially determines the order in which each traffic stream will be considered for grooming. Given a chromosome, they decoded the chromosome using a greedy heuristic, which follows a first-fit bin packing algorithm style. New offspring were generated using a $(\mu + \lambda)$ strategy, in which μ parents are used to generate λ children and all $(\mu + \lambda)$ solutions compete for the next generation.

7.3.5 Linear Programming Methods

Linear programming has also been employed to determine the optimal solution for the traffic grooming problem in SONET ring networks. In reference [22], the authors developed an integer linear programming (ILP) formulation for bidirectional WDM rings allowing arbitrary traffic between node pairs. They also presented approximate solutions using Simulated Annealing and a greedy approach. In reference [11], the authors also developed an ILP to minimize the total number of ADMs on a WDM ring network. Then, the complexity of the problem was reduced by relaxing a few integer variables to real variables but without compromising the solution quality for uniform traffic only. An extension to incorporate non-uniform traffic, using integer variables, was also provided. Finally, the authors also illustrated how this formulation can be extended to include dynamic traffic. Results were provided using fixed traffic only.

In [8], the authors formulated an ILP for the traffic grooming problem on unidirectional rings. The objective of the optimization problem was to minimize electronic routing. An interesting contribution of this work was the formulation of a sequence of bounds, where each successive bound in the sequence was tighter, or at least as good as the previous one. Each successive bound, however, involved more computation. The underlying principle in obtaining bounds was to decompose the ring into many sets of nodes, such

that in each set nodes were arranged in a path. A locally optimal solution was then employed within that set.

Recently, the use of SONET equipment equipped with transport blades operating at different line rates was suggested. The reason for doing this is to reduce the cost of implementation, since the cost of transport blades increases with the line rate. The optimal design of SONET rings with different line rates was studied using linear programming in a number of papers. In [12], an ILP was formulated to solve this problem. However, in [14] it was realized that an exact non-linear formulation can be decomposed into a number of smaller ILP formulations based on the fact that the identities of wavelengths being used at different rates are not as important as their number. Therefore, the ILP formulation corresponded to the number of combinations of wavelengths operating at different rates, which is significantly smaller than the number of wavelength permutations. This solution technique resulted in a significant saving in time and space.

7.3.6 Incremental Traffic

The following two studies considered incremental traffic patterns, and thus do not strictly fall under the category of static traffic. However, in both cases, once the initial static traffic has been accommodated, all of the new traffic are known at the same time and are treated collectively (as opposed to traffic arrival and accommodation separately). In [2], the authors defined a set of traffic matrices that reflects the changes in the traffic over a specified time period. The task was then to minimize the number of electronic ADMs needed to support any traffic matrix in this set; hence essentially supporting the traffic which dynamically changes within the given set of matrices. They considered a unidirectional ring. They also assumed that the traffic matrices are t -allowable, i.e., no node is sourcing more than t units of traffic at any time, where t is a specified constant. They then started out by placing an ADM on each wavelength at every node, and provided heuristics that remove the maximum possible number of ADMs, while supporting every t -allowable traffic matrix. They mentioned that their heuristic approaches can save up to 27% of the number of ADMs.

Similarly, researchers in [26] addressed the traffic grooming problem on unidirectional rings with incremental traffic using meta-heuristics. The authors assumed that they start with a given static traffic matrix and the network has been provisioned for that traffic matrix. Some time later, another traffic matrix is introduced. The objective is to accommodate the new traffic requests by disrupting as few current connections as possible. Two flavors of the problem are considered: best fit, in which they try to place as much new traffic as possible without adding any additional SONET ADMS, and full fit, in which they try to accommodate all the new traffic by adding minimum number of additional SONET ADMs. They presented an ILP, a

greedy heuristic, and a Tabu Search approach to solve the best fit case. They also presented an ILP and a Tabu Search based technique for the full fit case.

References

1. J.C. Bermond, D. Coudert (2003). Traffic grooming in unidirectional WDM ring networks: The all-to-all unitary case. In: Proceedings of Optical Network Design and Modeling, 2:1135–1153.
2. R. Berry, E. Modiano (2000). Reducing electronic multiplexing costs in SONET/WDM rings with dynamically changing traffic. IEEE J. Selected Areas in Communications, 18:1961–1971.
3. A.R. Billah, B. Wang, A.A. Awwal (2002). Effective traffic grooming in WDM rings. In: Proceedings of IEEE GLOBECOM, 3:2726–2730.
4. A. Chiu, E. Modiano (1998). Reducing electronic multiplexing costs in unidirectional SONET/WDM ring networks via efficient traffic grooming. In: Proceedings of IEEE GLOBECOM, 1:322–327.
5. A. Chiu, E. Modiano (2000). Traffic grooming algorithms for reducing electronic multiplexing costs in WDM ring networks. J. Lightwave Technology, 18:2–12.
6. W. Cho, J. Wang, B. Mukherjee (2001). Improved approaches for cost-effective traffic grooming in WDM ring networks: Uniform-traffic case. J. Photonic Network Communications, 3:245–254.
7. E.G. Coffman, M.R. Garey, D.S. Johnson (1997). Approximation algorithms for bin packing – A survey. In: Approximation Algorithms for NP-Hard Problems. PWS Publishing Company, Boston, pp. 46–93.
8. R. Dutta, G.N. Rouskas (2002). On optimal traffic grooming in WDM rings. IEEE J. Selected Areas in Communications, 20:110–121.
9. O. Gerstel, R. Ramaswami, G. Sasaki (1998). Cost-effective traffic grooming in WDM rings. In: Proceedings of IEEE INFOCOM, 1:69–77.
10. O. Gerstel, P. Lin, G. Sasaki (1999). Combined WDM and SONET design. In: Proceedings of IEEE INFOCOM, 2:734–743.
11. J.Q. Hu (2002). Traffic grooming in wavelength division multiplexing ring networks: A linear programming solution. J. Optical Networking, 1:397–408.
12. A. Jarray, B. Jaumard (2005). Exact ILP solution for the grooming problem in WDM ring networks. In: Proceedings of IEEE ICC 2005.
13. X.Y. Li, L.W. Liu, P.J. Wan, O. Frieder (2000). Practical traffic grooming scheme for single-hub SONET/WDM rings. In: Proceedings of 25th IEEE Conference on Local Computer Networks, pp. 556–564.
14. H. Liu, F. Tobagi (2004). A novel efficient technique for traffic grooming in WDM SONET with multiple line speeds. In: Proceedings of IEEE ICC 2004.
15. S.S. Roh, W.H. So, Y.C. Kim (2001). Design and performance evaluation of traffic grooming algorithms in WDM multi-ring networks. J. Photonic Network Communications, 3:335–348.
16. J.M. Simmons, E.L. Goldstein, A. Saleh (1998). On the value of wavelength-add/drop in WDM rings with uniform traffic. In: Proceedings of Optical Fiber Communication Conference, pp. 361–362.
17. J.M. Simmons, E.L. Goldstein, A. Saleh (1999). Quantifying the benefit of wavelength add-drop in WDM rings with distance-independent and dependent traffic. J. Lightwave Technology, 17:48–57.

18. J.M. Simmons, A. Saleh (1999). The value of optical bypass in reducing router size in gigabit networks. In: Proceedings of ICC, 1:591–596.
19. R. Ul-Mustafa, A.E. Kamal (2006). Grooming of non-uniform traffic on unidirectional and bidirectional rings. *Computer Communications*, 29:1065–1078.
20. P.J. Wan, L. Liu, O. Frieder (1999). Grooming of arbitrary traffic in SONET/WDM rings. In: Proceedings of IEEE GLOBECOM, 1B:1012–1016.
21. P.J. Wan, G. Calinescu, L. Liu, O. Frieder (2000). Grooming of arbitrary traffic in SONET/WDM BLSRs. *IEEE J. Selected Areas in Communications*, 18: 1995–2003.
22. J. Wang, W. Cho, V. Vemuri, B. Mukherjee (2001). Improved approaches for cost-effective traffic grooming in WDM ring networks: ILP formulations and single-hop and multi-hop connections. *J. Lightwave Technology*, 19:1645–1653.
23. Y. Xu, S.C. Xu, B.X. Wu (2002). Traffic grooming in unidirectional WDM ring networks using genetic algorithms. *Computer Communications*, 25:1185–1194.
24. X. Zhang, C. Qiao (1998). An effective and comprehensive approach to traffic grooming and wavelength assignment in SONET/WDM rings. In: Proceedings of SPIE Conference On All-optical Networking, 3531:221–232.
25. X. Zhang, C. Qiao (2000). An effective and comprehensive approach for traffic grooming and wavelength assignment in SONET/WDM rings. *IEEE/ACM Transactions on Networking*, 8:608–617.
26. S. Zhang, B. Ramamurthy (2003). Dynamic traffic grooming algorithms for reconfigurable SONET over WDM networks. *IEEE J. on Selected Areas in Communications*, 21:1165–1172.

Traffic Grooming in Next-Generation SONET/SDH Networks

Smita Rai, Biswanath Mukherjee

8.1 Introduction

Telecommunication networks are generally segmented into a three-tier hierarchy: access, metropolitan, and long-haul (and further delineations are also possible). Long-haul/backbone networks span inter-regional/global distances (1000 km or more in the US context, and perhaps an order of magnitude shorter in many other densely populated regions) and are optimized for transmission and related costs. At the other end of the hierarchy are access networks, providing connectivity to a plethora of customers within close proximity. Straddled in the middle are metropolitan (metro) networks, averaging regions between 10–100 km and interconnecting access and long-haul networks [1].

Metro networks today are based on synchronous digital hierarchy (SDH)/synchronous optical network (SONET) ring architectures (for a detailed reference, please see [2]). Namely, smaller tributary rings, e.g., OC-3/STM-1 (155 Mbps) or OC-12/STM-4 (622 Mbps), aggregate traffic onto larger core inter-office (IOF) rings that interconnect central office (CO) locations at higher bit rates, e.g., OC-48/STM-16 (2.5 Gbps). Overall, SONET/SDH has been very successful in delivering the early wave of end-user connectivity, namely voice. However, with the stupendous growth in Internet data traffic, new metro solutions are required to offer superior price/performance alternatives for legacy SONET/SDH expansion. For example, new platforms must offer high bandwidth scalability and carry multiple protocols over a common infrastructure to reduce costs. Various “data-centric” solutions are emerging, driven by advances in high-density electronic integrated circuit (IC) technology, and “next-generation” SONET/SDH is the most prominent.

Next-generation SONET¹ provides enhancements over SONET functionality that are pertinent to traffic grooming. We briefly discuss these features

¹ We mostly use the term SONET in the rest of this chapter; however the ideas are equally applicable to SDH.

below and point out the benefit from a grooming point of view, before discussing grooming approaches. See also Chapter 4 and its references for discussion of these features.

NG-SONET is comprised of the following technologies:

Generic Framing Procedure (GFP) [3]: Support for GFP makes it possible to transport any protocol over SONET, including Gigabit-Ethernet, Fiber Channel, etc. GFP is a traffic-adaptation protocol for broadband transport applications. It provides a standard mapping of either a physical layer or logical link layer signal to a byte-synchronous channel such as SONET/SDH links or wavelength channels in an optical transport network (OTN). There are two methods for mapping protocols onto GFP: frame-mapped GFP and transparent-mapped GFP. Frame-mapped GFP, optimized for a packet-switched environment, is the transport mode used for Point-to-Point Protocol (PPP), IP, and Ethernet traffic. Transparent-mapped GFP, intended for delay-sensitive storage-area network (SAN) applications, is the transport mode used for Fiber Channel (FC), Fiber Connection (FICON), and Enterprise Systems Connection (ESCON) traffic. Figure 8.1 (from [4]) shows the mapping of GFP to SONET/SDH using VC.

Virtual Concatenation (VC/VCAT) [5]: Virtual concatenation lets carriers allocate bandwidth within a SONET pipe to different services. For instance, the bandwidth can be allocated in 51.84 Mbps chunks (higher-order VCAT) based on the number of STS-1 channels used.

Link-Capacity Adjustment Scheme (LCAS) [6]: LCAS has been developed to allow the dynamic allocation of bandwidth to different services. It allows changes in bandwidth in response to customers requesting increases or decreases of bandwidth. LCAS, built on virtual concatenation, is a two-way signaling protocol that runs continuously between the source and destination of a bandwidth pipe. LCAS allows network operators to adjust the pipe capacity while it is in use (on the fly). It increases the possibility for on-demand traffic provisioning and on-line traffic grooming/re-grooming and makes SONET/SDH-based optical WDM network more data friendly.

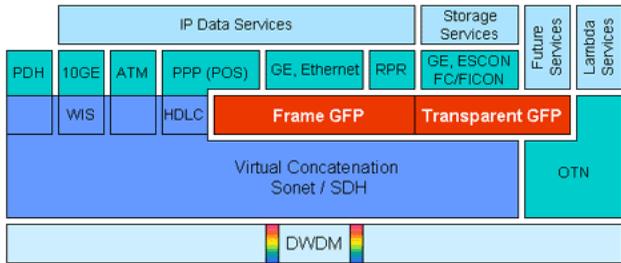


Fig. 8.1. GFP mapping to SONET/SDH.

VCAT, a feature that allows “inverse multiplexing” of traffic flows over multiple paths, has important consequences for traffic grooming, and we give a brief overview of this technology in the following section.

8.2 Virtual Concatenation

As the amount of data traffic rapidly increases, the inefficiency of transporting packet data through a SONET/SDH frame emerges as a major concern when network operators try to optimize the usage of their current bandwidth to support various types of new services and applications, based on IP, Frame Relay, Ethernet, Fiber Channel, etc.

In the traditional SONET/SDH multiplexing hierarchy, the frames of multiple low-speed traffic streams (say, STS-1 frame, approx. 51.84 Mbps) are combined to form the frame of a high-speed traffic stream. In order to support high-speed traffic from the same client source, e.g., a broadband ATM switch, N “contiguous” lower-order SONET/SDH payload containers are merged into one of greater capacity. This is called SONET/SDHs concatenation technique.

Usually, SONET/SDH’s concatenation is implemented at certain speeds, such as STS-3c, STS-12c, STS-48c, etc., which leads to a tiered bandwidth-allocation mechanism for different client services. Unfortunately, although “contiguous” and “tiered” concatenation is simple for implementation, it is not very flexible and not very efficient, especially in a multi-service network environment. From a single network node perspective, traffic streams from different client network equipment are to be discretely mapped into different tiers of SONET/SDH bandwidth trunks (data containers), which may result in large capacity wastage. For example, carrying a Gigabit Ethernet connection using a concatenated OC-48 pipe (approx. 2.5 Gbps) will lead to 60% bandwidth wastage.

From a network perspective, the time-slot contiguous requirement of a SONET/SDH concatenated channel imposes a constraint for traffic provisioning and may degrade network performance in a dynamic traffic environment where network resources are easy to be fragmented. The constraint also makes it more difficult for a network operator to perform efficient traffic grooming, i.e., packing different low-speed traffic streams onto high-capacity wavelength channels.

Virtual concatenation (VC or VCAT) will help a SONET/SDH-based optical network to carry data traffic in a finer granularity and hence use link capacity more efficiently. The basic principle of virtual concatenation [5] is quite simple. A number of smaller containers, which are not necessarily contiguous, are concatenated and assembled to create a bigger container that carries more data per second. Depending on a network’s switching granularity, virtual concatenation is possible for small container size from VT-1.5 up to STS-3c.

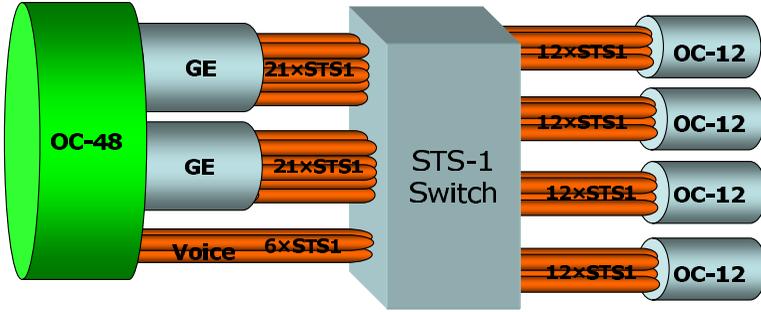


Fig. 8.2. An example of using VC to support different network services

Figure 8.2 (from [7]) shows an example of how to support multiple services using a single OC-48 channel through virtual concatenation. In Fig. 8.2, an OC-48 channel is used to carry two Gigabit Ethernet traffic streams and six STS-1 TDM voice traffic streams. Through a STS-1 switch, the traffic can be switched onto different OC-12 pipes, and these OC-12 pipes can be sent through the network over various routes. Figure 8.2 also illustrates the potential load-balancing benefit that VC can provide to a transport network.

One important thing to note is that, when multiple traffic streams from one client are sent over different routes, the VC mappers at the destination nodes need to compensate for the differential delay between the bifurcated streams when they are reconstructed at the destination node. Currently, such a typical, commercially available device may support up to 50 ms (+/-25 ms) delay with external RAM, which is equivalent to a 10,000 km difference in route length [8]. In general, a virtually concatenated SONET/SDH channel made up of $N \times$ STS-1 is transported as individual STS-1s across the network; and, at the receiver, the individual STS-1s are re-aligned and sorted to recreate the original payload. Figure 8.3 (from [1]) shows an overview of a multi-service SONET/SDH-based optical network employing VC technology. VC can be supported at the edge optical crossconnects (OXC) (in port cards) or in separate traffic-aggregation network elements connecting client network equipment and OXC.

8.2.1 Benefits of Virtual Concatenation

From a network perspective, the SONET/SDH VC technique can provide the following benefits to an optical network:

1. *Relax time-slot alignment and continuity constraints.* Instead of aligning to particular time-slots and consisting of N contiguous STS-1 time-slots within a wavelength, a high-speed STS- N channel could be constructed from any N STS-1 time-slots and carried by different wavelength channels.

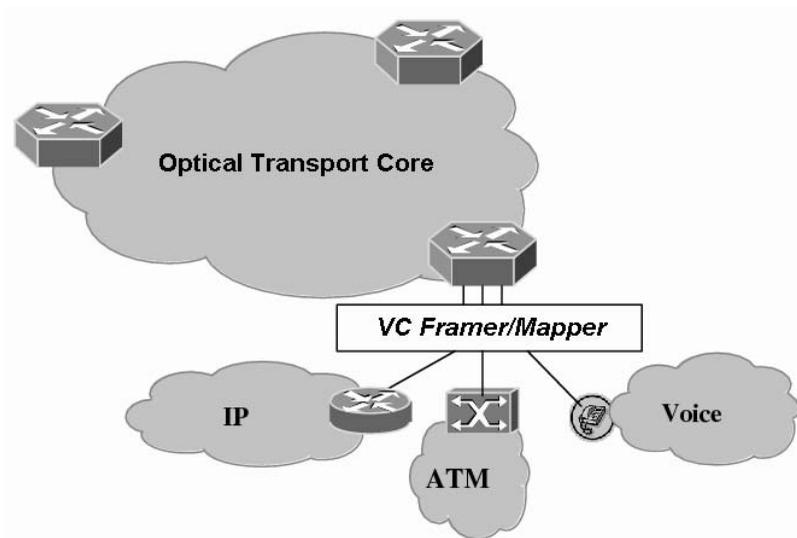


Fig. 8.3. An overview of a VC-enabled multi-service optical network

2. *More efficiently utilize channel capacity to support multiple types of data and voice services.* Instead of mapping data traffic (packet/cell/frame) into SONET/SDH frames in a discrete tiered manner, optical networks can carry data traffic in a more resource-efficient way. Traffic granularity can be increased in the unit of 1.6 Mbps (VT-1.5) in a metro-area network, and 48 Mbps (STS-1) or 150 Mbps (STS-3c) in an optical backbone network.
3. *Bifurcate traffic streams to balance network load.* With VC, it is possible to split a high-speed traffic stream into multiple low-speed streams and route them separately through the network. This enables the traffic to be distributed across the network more evenly; and, hence, it can improve the network's blocking performance.

8.3 Algorithm for Provisioning with Virtual Concatenation

The benefits of SONET/SDH VC to an optical WDM mesh network (shown in Fig. 8.4) under dynamic traffic environment is investigated via simulations in [9]. Connections with different bandwidth granularities are assumed to come and leave the network, one at a time, following a Poisson arrival process and negative-exponential-distribution holding time.

To distinguish the benefit effects, two types of traffic pattern are studied (pattern I and II), one consisting of five service classes and the other having ten service classes. Capacity of each wavelength channel is assumed to be OC-192. In traffic pattern I, data rates for each service class are approximately 51

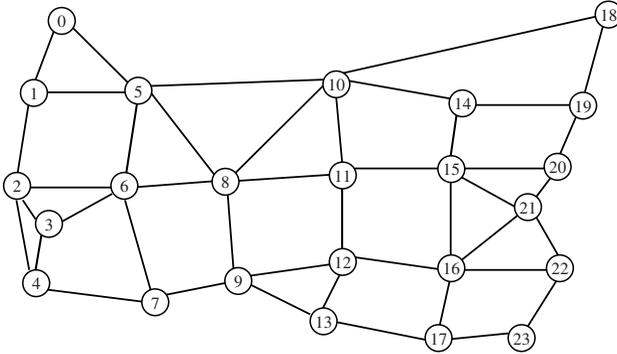


Fig. 8.4. A 24-node example network topology

Mbps, 153 Mbps, 622 Mbps, 2.5 Gbps, and 10 Gbps, which can be perfectly mapped into the tiered SONET/SDH containers and their corresponding optical carriers, i.e., OC-1, OC-3c, OC-12c, OC-48c, and OC-192c. The service characteristic of pattern II, i.e., service classes, service rate, and corresponding SONET/SDH containers with or without VC is shown in Table 8.1.

When traffic bifurcation is needed, a simple route-computation and traffic-bifurcation heuristic is applied to a connection request. The problem of finding the set of minimal-cost routes with enough aggregated capacity for a request between a given node pair in a network is NP-Complete. The heuristic works as follows:

- Step 1: A shortest path is computed according to network administrative cost between the node pair.
- Step 2: The bandwidth of the route is calculated. The bandwidth of the route is constrained by the link along the route, which has minimal free capacity. Then, update the available capacity of the links along the route (i.e., decrease the available capacity by the minimal capacity).
- Step 3: Remove the link without free capacity and repeat Steps 1 and 2 until the connection can be carried by the set of routes computed or no more routes exist for the connection.

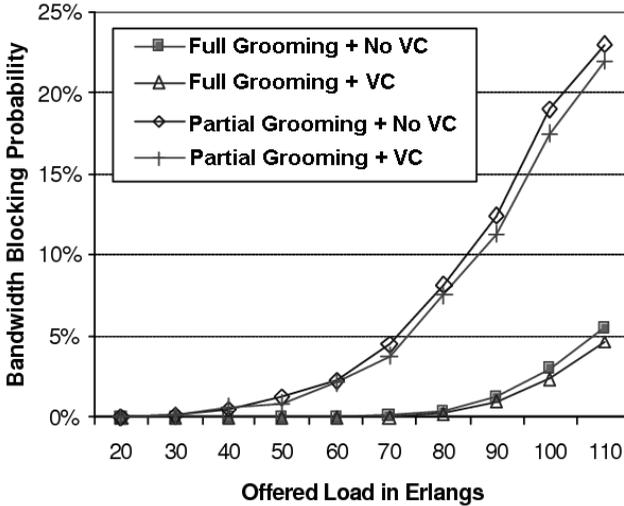
Note that, depending on the implementation, the VC mappers at the receiver end nodes may only be able to handle a limited number of routes

Table 8.1. Traffic pattern II: ten service classes

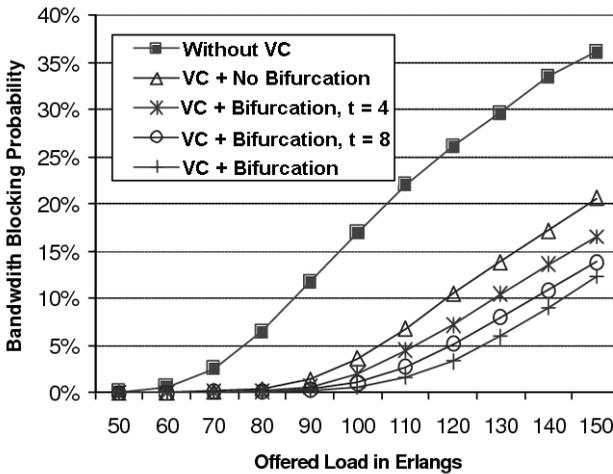
Class	Rate	Without VC	With VC	Class	Rate	Without VC	With VC
1	50 M	STS-1	STS-1	6	600 M	STS-12c	STS-12
2	100 M	STS-3c	STS-2	7	1 G	STS-48c	STS-21
3	150 M	STS-3c	STS-3	8	2.5 G	STS-48c	STS-48
4	200 M	STS-12c	STS-4	9	5 G	STS-192c	STS-96
5	400 M	STS-12c	STS-8	10	10 G	STS-192c	STS-192

(denoted by t) for a single connection, in which case, the connection will be blocked after t routes have been examined.

Figure 8.5 illustrates network performance in terms of bandwidth blocking probability (BBP) as a function of network offered load in Erlangs. BBP is considered as the measurement metric because connections from different service classes may have different bandwidth requirements. Network offered load



(a) Five service classes.



(b) Ten service classes.

Fig. 8.5. Illustrative numerical results

is normalized to the unit of OC-192. A 24-node, 43-bidirectional-link network topology is used in our study (see Fig. 8.4). Each link has 16 wavelengths.

Figure 8.5(a) shows the network performance for traffic pattern I with or without employing VC. Two types of network configurations are examined, i.e., all nodes are either equipped with STS-1 full-grooming switches or partial-grooming switches. Note that, in a partial-grooming switch, only a limited number of wavelength channels (6 in our simulation) can be switched to a separate grooming switch (or grooming fabric within an OXC) to perform traffic grooming.

As one can observe from Fig. 8.5(a), there is around 5–10 percent network performance gain through VC technique. In traffic pattern I, every service class can be perfectly mapped into one of the tiered SONET/SDH containers, and we assumed that no traffic bifurcation is allowed. Therefore, the performance improvement shown in Fig. 8.5(a) comes solely from the capability of eliminating the time-slot alignment and contiguity constraints provided by VC to the network.

Figure 8.5(b) shows how VC can significantly improve network performance when the network needs to support data-oriented services with different bandwidth requirements under the network configuration in which full-grooming OXCs are used at every node. It is straightforward to see that BBP is significantly reduced by employing VC. Meanwhile, more improvement can be achieved by allowing a simple traffic-bifurcation scheme. Note that, in the study, a connection will be bifurcated if and only if no single route with enough capacity exists. The results from different values of t have been examined and some of them are shown in Fig. 8.5(b), i.e., 4, 8, and unlimited. It can be expected that more advanced network load-balancing and traffic-bifurcation approaches can further optimize network throughput.

The benefits quantitatively demonstrated above arise from more efficient grooming that exploits the flexibility provided by Virtual Concatenation (VC). Since a connection can be split up, and routed across multiple paths, and there are fewer constraints on time-slot alignment, grooming techniques that pack connections more effectively can be developed. SONET/SDH VC could also benefit an optical transport network on other aspects, such as network compatibility, network resiliency, network management and control, etc. VC works across legacy networks. Only the end nodes of the network are aware of the containers being virtually concatenated. In terms of resiliency, since individual members of a virtually concatenated channel may be carried through different routes, a network failure may only affect partial bandwidth of a connection service. In this case, an unprotected, best-effort connection may still get the service under the reduced bandwidth before the network failure is fixed, and a high-priority connection can still be provided partial service before the network protection/restoration operation is active to restore the full service. Reference [10] explores reliability-based provisioning in next-generation SONET/SDH networks addressing some of these concepts of degraded service and expected bandwidth for a multi-path connection.

Reference [11] investigates novel schemes for protecting multi-path connections by introducing concepts such as “intra-connection” sharing apart from the conventional “inter-connection” sharing.

8.4 Next-Generation SONET in Interconnected Rings

In this section, we explore the problem of provisioning connections in dual-node-interconnected SONET/SDH ring networks. Dual-node interconnection using drop-and-continue architecture (for more details, please refer to [12, 13]) is the de-facto SONET standard for connecting rings – Bi-directional Line Switched Rings (BLSR) or Uni-directional Path Switched Rings (UPSR) – to ensure survivability against all single-link and single-node failures. Similar recommendations exist for SDH networks [13]. This architecture, combined with legacy SONET time-slot assignment rules for contiguous-concatenated signals (equivalent rules exist for SDH), places some unique constraints on provisioning connections dynamically on such ring networks. We explore the benefits of employing virtual concatenation and grooming high-bandwidth connections over multiple paths in such networks.

We developed an algorithm called `PATH.FINDER` [14] to find valid routes in such a network of interconnected rings. The basic intuition behind the algorithm is as follows. Within a ring, there are not too many choices of alternate paths. For an intra-ring demand, we can only choose one of two directions in which the connection should travel for BLSR rings. However, when an inter-ring connection has to travel through several intermediate rings, the choice of the next-hop ring becomes crucial, at each stage. We capture this intuition by constructing a simplified graph from the current state of the network.

This graph contains the source and destination of the connection request, as well as all the interconnection pairs between each ring. These dual-node interconnection pairs on a ring are collapsed into one node in the simplified graph (see Fig. 8.6). We add an edge between two nodes in the simplified graph, called an auxiliary graph, if there exists a valid path between them in the original network (see Fig. 8.7).

Note that the above construction is not necessary for demands whose source and destination nodes are within the same ring. For BLSR rings, the direction in which to send the demand can be chosen based on some path metric.

We also developed an intelligent cost and time-slot assignment algorithm (called `SmartCost`) to minimize fragmentation while provisioning a connection. For details, the reader is referred to [14].

To study the properties of our provisioning algorithm, we simulated a dynamic network environment. The network topology is shown in Fig. 8.8. It is composed of 8 OC-192 BLSR rings with a total of 50 nodes and 12 interconnection pairs. The large number of interconnection pairs reduces the effect of the drop-and-continue penalty, by allowing traffic to be load balanced

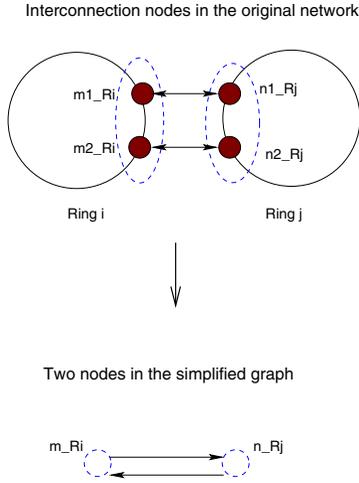


Fig. 8.6. Graph transformation

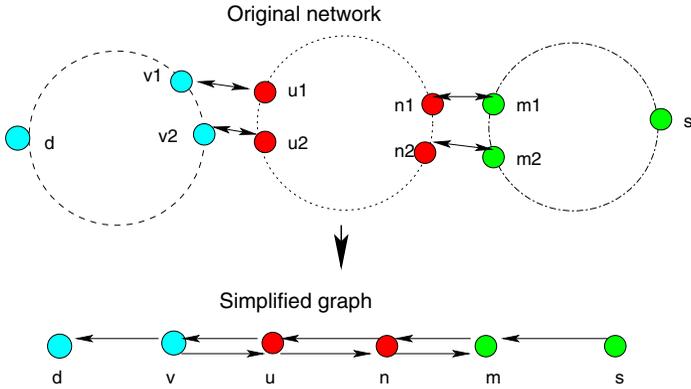


Fig. 8.7. Auxiliary graph

over different interconnection pairs, as was suggested in [15]. The connection-arrival process is Poisson and the connection-holding time follows a negative exponential distribution, in which source and destination nodes are selected uniformly from among the nodes in the network with inter-ring demands and intra-ring demands being in the ratio 25:75. The traffic mix is as follows: $STS-1:STS-3c:STS-12c:STS-48c = 100:50:10:1$. This is based on the assumption that lower-bandwidth requests come in at a larger proportion than higher-bandwidth requests. All connections are bi-directional. Note that the contiguous concatenation and alignment constraints are too restrictive for STS-48c in the network topology using OC-192 rings. Such a request can

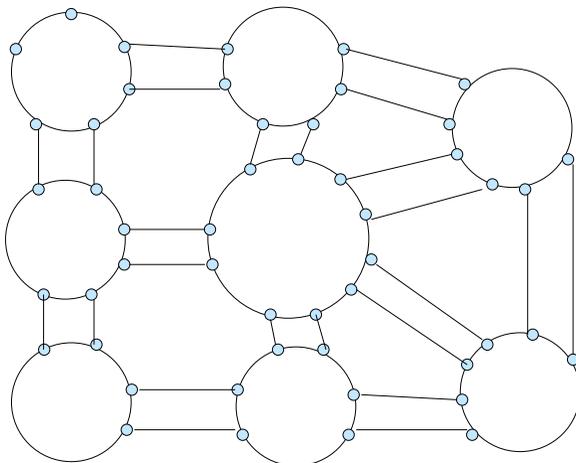


Fig. 8.8. Ring topology

use slots starting at only two time slots, namely 1 and 49 on any link, and all the adjacent 47 slots must also be free.

We compared our SmartCost algorithm to a shortest-hop approach with first-fit to provision paths. In first-fit time-slot assignment, the connection is assigned to the first free set of time slots which satisfy the contiguity and alignment requirements. Both algorithms use our graph abstraction (our PATH_FINDER module) to select a route out of a set of valid routes; the differences lie in how cost is assigned to a path and how time slots are assigned.

As a measure of efficient capacity utilization, we compare the blocking ratio (also called blocking probability) which is the ratio of the number of connections blocked for a particular granularity to the number of requests for that granularity, as well as the overall bandwidth blocking ratio. For all our simulation results, we find the 95% confidence intervals according to the t -distribution [16] by performing several trials of each experiment.

We observed a high blocking ratio of STS-48 with both the algorithms (Fig. 8.9), true to our intuition that contiguous concatenation imposes significant penalty on higher-bandwidth requests. We explored how next-generation SONET technologies, such as VCAT, can help to significantly improve the blocking ratio of these high-bandwidth requests by grooming them over multiple paths. From our simulation results in Fig. 8.9, we observed that a significant chunk of high-bandwidth connections, such as STS-48, are blocked without the facility of VCAT. We incorporated virtual concatenation in our algorithms, allowing an STS-48 signal to be broken up into 48 STS-1 signals, and routed independently in the network. The other connections are carried using contiguous concatenation as before. We observe substantial improvement in the blocking ratio of STS-48 requests (Fig. 8.9), and significant improvement in capacity utilization. The total bandwidth blocking ratio

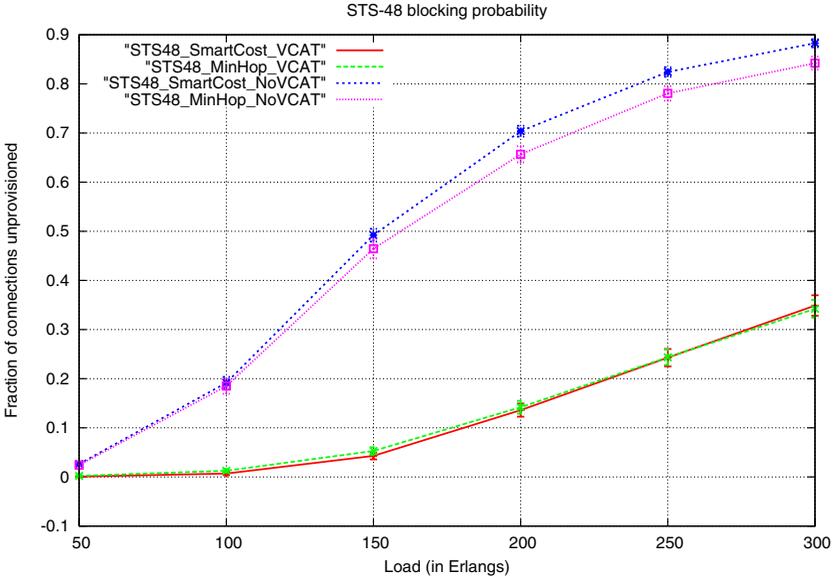


Fig. 8.9. STS-48 blocking ratio with and without VCAT

with the introduction of VCAT decreases to 1/3 of the value without it at intermediate loads (Fig. 8.10). There is a slight increase in the fraction of STS-12c connections being provisioned (Fig. 8.11), since capacity is used up by STS-48 bandwidth connections; however the overall bandwidth blocking

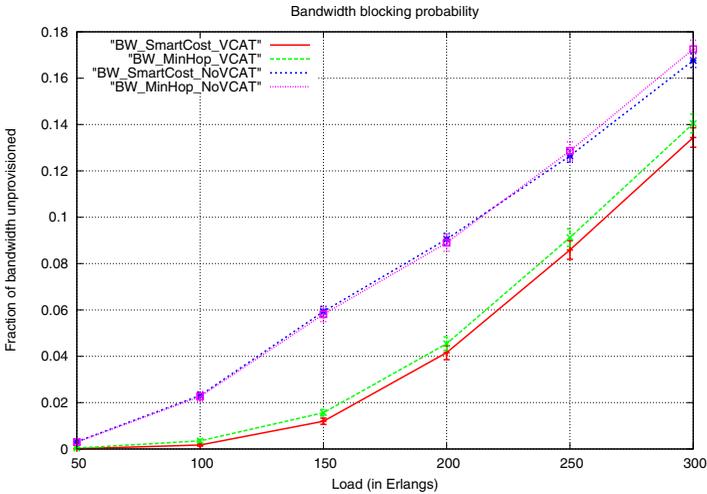


Fig. 8.10. Bandwidth blocking ratio with and without VCAT

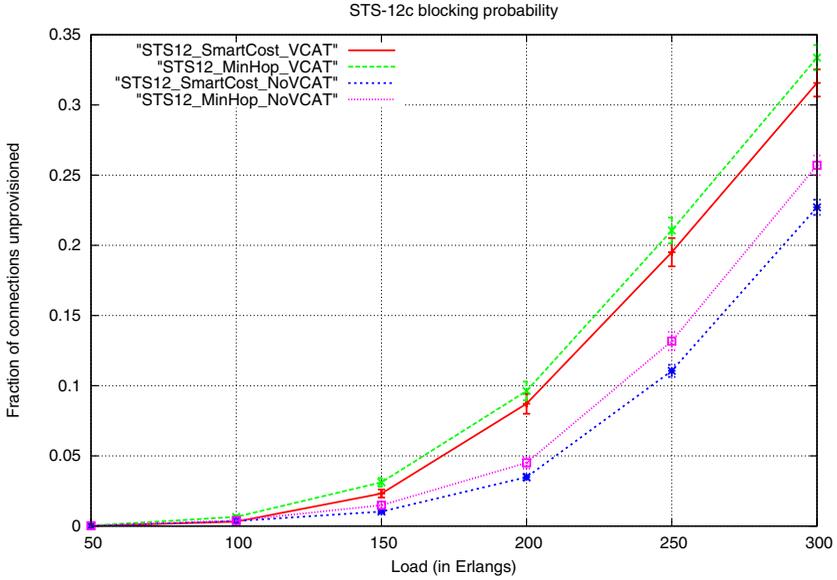


Fig. 8.11. STS-12c blocking ratio with and without VCAT

ratio improves significantly. The blocking ratio of lower-bandwidth requests shows negligible difference (Fig. 8.12), and for STS-1 the ratio is effectively zero for all loads and all strategies. SmartCost algorithm performs better than the minimum-hop scheme; however, substantial improvement due to flexible

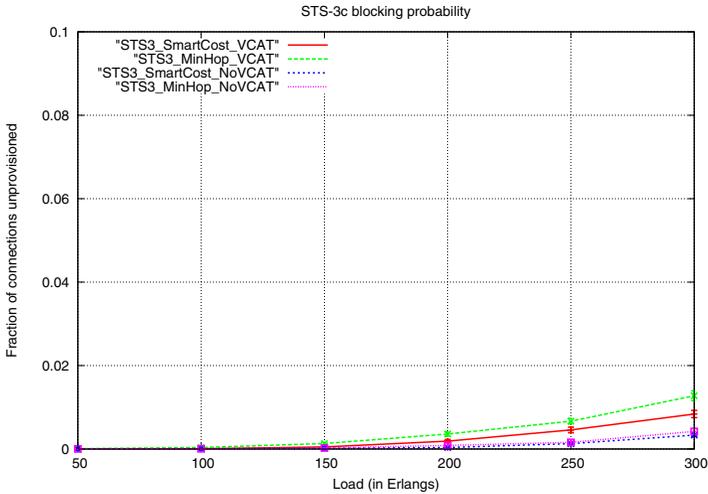


Fig. 8.12. STS-3c blocking ratio with and without VCAT

concatenation is observed when comparing any approach without VCAT to a VCAT-enabled approach.

With the ability to split up and route a connection over different paths, there arises the problem of differential delay among the different paths. The standards allow a delay up to 256 ms, but practical values range from 60 to 100 ms [17]. Addressing differential delay issues are out of scope of this chapter; the interested reader is referred to [18, 19] for some work on differential-delay-based provisioning. Note that, since we focus on a network of rings, if the STS-48c demands are split up using VCAT, for an intra-ring demand, it is expected that the differential delay would not pose a serious problem (there are only two directions to choose from and a BLSR ring can have a maximum 16 nodes, with 1200 km of total fiber [2]). For an inter-ring demand, in our algorithms, we do not allow paths to loop around the same ring multiple times. For instance, paths such as $Ring_A \rightarrow Ring_B \rightarrow Ring_C \rightarrow Ring_B$ are not permitted. Due to the limited number of rings, and limits on the maximum number of nodes on a ring as well as the total fiber, we expect that differential delay issues would be less problematic. Further exploration of delay issues is an open problem for future research.

8.5 Conclusion

Virtual concatenation (VC / VCAT) and link-capacity adjustment scheme (LCAS) can help a SONET/SDH-based optical network to evolve towards a data-centric intelligent automatically switched optical network. They offer the following benefits: (a) relaxing the time-slot continuity and alignment constraints of traditional SONET/SDH concatenation; (b) improving bandwidth efficiency of a wavelength channel; (c) enabling inverse-multiplexing (i.e., traffic bifurcation) and load balancing; and (d) improving service resilience. By loosening the constraints on time-slot assignment and by allowing more flexibility in provisioning, these benefits allow for the development of efficient traffic grooming techniques that can pack connections more effectively in SONET/SDH pipes. We studied these benefits in a general mesh network as well as in a network of interconnected rings using dual-node interconnection. We presented illustrative numerical results to demonstrate the significant benefits which can be obtained by employing virtual concatenation.

Our simple provisioning algorithms that groom connections onto multiple paths, exploiting the inverse multiplexing capability of VCAT, show significant performance gains, when compared to approaches using legacy SONET/SDH.

Moreover, built on virtual concatenation, LCAS allows a network operator to adjust the pipe capacity while it is in use (on the fly). This increases the possibility for on-demand traffic provisioning and it makes SONET/SDH-based optical WDM networks more data friendly.

References

1. B. Mukherjee, *Optical WDM Networks*, Springer, Optical Network Series, 2006.
2. W. Goralski, *SONET/SDH*, 3rd edition, McGraw Hill, 2002.
3. ITU-T, Recommendation G.704/Y.1303, "Generic Framing Procedure (GFP)," 2001.
4. S. Stanley, "Making SONET Ethernet-friendly," Lightreading report, March 2003.
5. ITU-T, Recommendation G. 707, "Network node interface for the synchronous digital hierarchy (SDH)," April 2002.
6. ITU-T, Recommendation G.7042/Y.1305, "Link capacity adjustment scheme (LCAS) for virtual concatenated signals," 2001.
7. S. Stanley, "Next-gen SONET silicon," Lightreading report, June 2002.
8. T. Hills, "Next-gen SONET," Lightreading report, May 2002.
9. K. Zhu, J. Zhang, and B Mukherjee, "Ethernet-over-SONET (EoS) over WDM in Optical Wide-Area Networks (WANs): Benefits and Challenges," *Photonics Network Communications*, vol. 10, no. 1, pp. 107–118, July 2005.
10. S. Rai, O. Deshpande, C. Ou, and B. Mukherjee, "Reliable multi-path provisioning for high-capacity optical backbone mesh networks," *Proc., IEEE International Conference on Communications*, June 2005. Extended version to appear in *IEEE/ACM Transactions on Networking*, vol. 15, no. 4, August 2007.
11. C. Ou, L. H. Sahasrabudde, K. Zhu, C. U. Martel, and B. Mukherjee, "Survivable virtual concatenation for data over SONET/SDH in optical transport networks," *IEEE/ACM Transactions on Networking*, vol. 14, no. 1, Feb. 2006.
12. G. W. Ester, "SONET: Interconnecting rings using drop and continue," *Telephone Engineer and Management*, June 1, 1993.
13. ITU-T Recommendation, "Interworking of SDH network protection architectures," G.842, April 1997.
14. S. Rai, L. Song, C. -F. Su, T. Hamada, and B. Mukherjee, "A novel provisioning framework for dual-node interconnected SONET/SDH rings," *Proc., Optical Fiber Communications*, March 2006.
15. B. T. Doshi, S. Dravida, P. Harshavardhana, P. K. Johri, and R. Nagarajan, "Dual (SONET) ring interworking: High penalty cases and how to avoid them," *Proc., International Teletraffic Congress ITC 15*, Elsevier 1997.
16. K. S. Trivedi, *Probability and Statistics with Reliability, Queuing, and Computer Science Applications*, Englewood Cliffs, NJ: Prentice-Hall, 1982.
17. G. Bernstein, B. Rajagopala, and D. Saha, *Optical Network Control: Architecture, Protocols and Standards*, Addison-Wesley, July 2003.
18. A. Srivastava, S. Acharya, M. Alicherry, B. Gupta, and P. Risbood, "Differential Delay aware Routing for Ethernet over SONET/SDH," *Proc., INFOCOM'05*, March 2005.
19. S. Huang, S. Rai, and B. Mukherjee, "Survivable differential delay aware multi-service over SONET/SDH networks with virtual concatenation," *Proc., Optical Fiber Communications (OFC) Conference*, March 2007.

Mathematical Programming Approaches

JianQiang Hu

9.1 Introduction

There are two basic architectures used in WDM networks: ring and mesh. The majority of optical networks in operation today have been built based on the ring architecture. However, carriers have increasingly considered the mesh architecture as an alternative for building their next generation networks. Various studies have shown that mesh networks have a compelling cost advantage over ring networks. Mesh networks are more resilient to various network failures and also more flexible in accommodating changes in traffic demands (e.g., see [4, 6, 17] and references therein). In order to capitalize on these advantages, effective design methodologies are required.

The grooming problem in an optical mesh network in fact involves traffic grooming, routing, and wavelength assignment. The problem of traffic grooming and routing (GR) for mesh networks is to determine how to efficiently route traffic demands and at the same time to combine lower-rate (sub-wavelength) traffic demands onto a single wavelength. On the other hand, the problem of wavelength assignment (WA) is to determine how to assign specific wavelengths to lightpaths, usually under the wavelength continuity constraint. In previous studies on the routing and wavelength assignment (RWA) problem for mesh networks (e.g., see [12, Chapter 8] and references therein), the issue of traffic grooming has largely been ignored, i.e., it has been assumed that each traffic demand takes up an entire wavelength. In practice, this is hardly the case, and networks are typically required to carry a large number of lower-rate traffic demands. We should point out that the grooming problem is an NP-complete problem; therefore, heuristics and approximations are usually needed for solving this problem.

The traffic grooming problem for mesh networks has only been considered recently (e.g., see [5, 7, 8, 10, 16]). The objective considered in [16] is either to maximize the network throughput or to minimize the connection-blocking probability, which are operational network-design problems. A different operational network-design problem is considered in [10], in which a Lagrangian

relaxation based method is proposed to minimize the total cost associated with links. Alternatively, a strategic network-design problem is to minimize the total network cost. Typically, the cost of a nation-wide optical network is dominated by optical transponders and optical amplifiers. If one assumes that the fiber routes are fixed (e.g., see [14, 15] for some recent work on the problem of amplifier allocation), then the amplifier cost is constant, in which case one should concentrate on minimizing the number of transponders in the network. Ding and Hamdi [5] proposed a heuristic algorithm to minimize the number of transponders as well as the number of wavelengths in mesh networks. In [8], Lee and Park proposed a genetic algorithm to minimize a combination of the number of transponders and the number of wavelengths. In [7], Hu and Leida first formulated the grooming problem as an integer linear programming (ILP) problem, then decomposed it into two smaller problems: the GR problem and the WA problem.

This chapter is mainly based on the work of [7], where the problem of traffic grooming, routing, and wavelength assignment (GRWA) was considered with the objective of minimizing the number of transponders in the network. We first formulate the GRWA problem as an integer linear programming (ILP) problem. Unfortunately, the resulting ILP problem is usually very hard to solve computationally, in particular for large networks. To overcome this difficulty, we then propose a decomposition method that divides the GRWA problem into two smaller problems: the traffic grooming and routing (GR) problem and the wavelength assignment (WA) problem. In the GR problem, we only consider how to groom and route traffic demands onto lightpaths (with the same objective of minimizing the number of transponders) and ignore the issue of how to assign specific wavelengths to lightpaths. Similar to the GRWA problem, we can formulate the GR problem as an ILP problem. The size of the GR ILP problem is much smaller than its corresponding GRWA ILP problem. Furthermore, we can significantly improve the computational efficiency for the GR ILP problem by relaxing some of its integer constraints, which usually leads to quite good approximate solutions for the GR problem. Once we solve the GR problem, we can then consider the WA problem, in which our goal is to derive a feasible wavelength assignment solution.

We note that the WA problem has been studied by several researchers before (e.g., see [1, 2, 3, 9, 11, 12, 13] and references therein). However, the objective in all these studies has been to minimize the number of wavelengths required in a network, in some cases by using wavelength converters. In general, the use of additional wavelengths in a network only marginally increases the overall network cost as long as the total number of wavelengths used in the network does not exceed a given threshold (the wavelength capacity of a WDM system). This is mainly because the amplification cost is independent of the number of wavelengths. In recent years, the wavelength capacity for optical networks has increased dramatically. For example, with most advanced techniques, a single WDM system on a pair of fibers can carry up to 160

10G-wavelengths or 80 40G-wavelengths. Of course, once the wavelength capacity is exceeded, then a second parallel system (with another set of optical amplifiers) needs to be built, which would then substantially increase the network cost. Therefore, assuming a single WDM system on all fiber routes fixes the amplifier cost, then one should focus on minimizing the number of transponders in the network, which is already taken into consideration in the GR problem. In this setting, the objective in our WA problem is to find a feasible wavelength assignment solution under the wavelength capacity constraint.

It is clear that, in general, the decomposition method would not yield the optimal solution for the GRWA problem. However, we will provide a sufficient condition under which we show that the decomposition method does produce an optimal solution for the GRWA problem. This is achieved by developing a simple algorithm that, under this sufficient condition, finds an optimal wavelength assignment.

The rest of this chapter is organized as follows. In Section 9.2, we present the GRWA problem and demonstrate how it can be formulated as an ILP problem. In Section 9.3, we first present our decomposition method. We then provide an ILP formulation for the GR problem and develop an algorithm for solving the WA problem. We also discuss under what condition the decomposition method produces an optimal solution for the GRWA problem. Some numerical results are provided in Section 9.4.

9.2 The GRWA Problem

An optical mesh network architecturally has two layers: a physical layer and an optical layer. The physical layer consists of fiber spans and nodes and the optical layer consists of lightpaths (optical links) and a subset of nodes contained in the physical layer. A lightpath in the optical layer is a path connecting a pair of nodes via a set of fiber spans in the physical layer. Throughout this chapter, we assume that lightpaths and their routes in the physical layer are given. In practice, the selection of lightpaths is another important design issue that needs to be addressed, which is beyond the scope of this chapter.

We use graph $G_f = (V_f, E)$ to represent the physical layer, where E is the set of edges representing fiber spans and V_f is the set of nodes representing locations which are connected via fiber spans. We use graph $G_o = (V_o, L)$ to represent the optical layer, where L is the set of edges representing lightpaths and $V_o \subset V_f$ is a subset of locations that are connected via lightpaths. Each edge in L corresponds to a path in G_f . Here we treat each lightpath as a logical connection between a pair of nodes (not just a single wavelength); therefore, one lightpath can contain multiple wavelengths. For ease of exposition, we first assume that G_o is a directed graph (i.e., the lightpaths are unidirectional). The extension to the undirected graph case is quite straightforward

and will be discussed later in this section (basically, we can simply replace every undirected edge with two directed edges).

The GRWA problem studied in this chapter can be described as follows. Assuming that a set of traffic demands are given (some of them are of low rate, i.e., sub-wavelength), our goal is to find an optimal way to route and groom these demands in the optical layer, G_o , and also to assign a set of specific wavelengths to each lightpath so that the total number of transponders required is minimized. There are two key constraints we need to take into consideration in this problem: (1) the wavelength capacity constraint for each fiber span, and (2) the wavelength continuity constraint for every lightpath, i.e., the same wavelength(s) needs to be assigned to a lightpath over the fiber spans it traverses (we assume that wavelength converters are not used; therefore the wavelength continuity constraint is required). In this problem setting, the number of transponders required for each lightpath is equal to twice the number of wavelengths assigned to it (one transponder for each end of each wavelength on a lightpath). Therefore, by grooming several low rate demands onto a single wavelength, we can potentially reduce the total number of wavelengths required by the lightpaths, thus the number of transponders.

The GRWA problem can be formulated as an integer linear programming (ILP) problem. First, we need to introduce some necessary notation:

W : the set of wavelengths available on each fiber;

D : the set of traffic demands;

g : the capacity of a single wavelength;

s_d : the size of demand $d \in D$;

A : $= [a_{v,l}]_{|V_o| \times |L|}$, the node-edge incidence matrix of graph G_o , where $a_{v,l} = 1$ if lightpath l originates from node v , -1 if lightpath l terminates at node v , and 0 otherwise;

B : $= [b_{e,l}]_{|E| \times |L|}$, the fiber-lightpath incidence matrix, where $b_{e,l} = 1$ if fiber span e is on lightpath l , and 0 otherwise;

u_d : $= [u_{v,d}]_{v \in V_o}$, the source-destination column vector for $d \in D$, where $u_{v,d} = 1$ if v is the starting node of d , -1 if v is the end node of d , and 0 otherwise;

x_d : $= [x_{l,d}]_{l \in L}$, the column vector containing lightpath routing variables for $d \in D$, where $x_{l,d} = 1$ if demand d traverses lightpath l , and 0 otherwise;

y_w : $= [y_{l,w}]_{l \in L}$, the column vector containing wavelength assignment variables for $w \in W$, where $y_{l,w} = 1$ if wavelength w is assigned to lightpath l , and 0 otherwise (note that in our setting each lightpath l is treated as a logical connection between a pair of nodes, hence it can be assigned with multiple wavelengths, i.e., it is possible that $\sum_{w \in W} y_{l,w} \geq 1$);

$\mathbf{1}$: $= [1, 1, \dots, 1]$, the unit column vector of appropriate size.

Then the GRWA problem can be formulated as the following ILP problem (which we shall refer to as the GRWA ILP problem):

$$\min \sum_{w \in W, l \in L} y_{l,w}$$

$$\text{s.t. } Ax_d = u_d \quad d \in D \quad (9.1)$$

$$By_w \leq \mathbf{1} \quad w \in W \quad (9.2)$$

$$\sum_{d \in D} s_d x_{l,d} \leq g \sum_{w \in W} y_{l,w} \quad l \in L \quad (9.3)$$

x and y are binary variables.

where the objective function $\sum_{w \in W, l \in L} y_{l,w}$ is the total number of wavelengths assigned to all lightpaths, which is equivalent to minimizing the total number of transponders needed. The three constraints are:

- Equation (9.1) is the flow balance equation, which guarantees that the lightpaths selected based on x_d constitute a path from the starting node of d to the end node of d .
- Equation (9.2) implies a single wavelength along each fiber span can be assigned to no more than one lightpath.
- Equation (9.3) is the capacity constraint for lightpath l , since $\sum_{d \in D} s_d x_{l,d}$ is the total amount of demands carried by lightpath l and $g \sum_{w \in W} y_{l,w}$ is the total capacity of lightpath l .

We refer to the type of the network considered above as the basic model. There are several variations of the basic model, which include:

1. networks with both protected and unprotected demands;
2. networks in which lightpaths are undirected;
3. networks with non-homogeneous fibers where different types of fiber may have different wavelength capacities; and
4. networks in which demand exceeds a single WDM system per fiber pair.

9.3 A Decomposition Method

In the previous section, we formulated the GRWA problem as an ILP problem; however, it may not be computationally feasible to solve the ILP problem, particularly for large networks (e.g., see numerical results in Section 9.4). Therefore, it is necessary to find more efficient ways to solve the GRWA problem. In this section, we propose a decomposition method that divides the GRWA problem into two smaller problems: the traffic grooming and routing (GR) problem and the wavelength assignment (WA) problem. In the GR problem, we only consider how to groom and route demands over lightpaths and ignore the issue of how to assign specific wavelengths to lightpaths. Based on the grooming and routing, we can then derive wavelength capacity requirements for all lightpaths. Similar to the GRWA problem, we formulate the GR problem as an ILP problem. The size of the GR ILP problem is much smaller than its corresponding GRWA ILP problem. Furthermore, we can

significantly improve the computational efficiency for the GR ILP problem by relaxing some of its integer constraints, which usually leads to approximate solutions for the GR problem. Once we solve the GR problem, we can then consider the WA problem, in which our goal is to derive a feasible wavelength assignment solution that assigns specific wavelengths to lightpaths based on their capacity requirements derived in the GR problem.

It is obvious that, in general, the decomposition method would not yield the optimal solution for the GRWA problem. However, we will provide a sufficient condition under which we show that the decomposition method does produce an optimal solution for the GRWA problem. We also develop a simple algorithm that finds a wavelength assignment solution under this sufficient condition.

9.3.1 The GR Problem

Let $t = [t_l]_{l \in L}$, a column vector containing lightpath capacity decision variables, where $t_l = \sum_{w \in W} y_{l,w}$ is the number of wavelengths needed for lightpath $l \in L$. Then, the GR problem can be formulated as:

$$\begin{aligned} \min \quad & \sum_{l \in L} t_l \\ \text{s.t.} \quad & Ax_d = u_d \quad d \in D \end{aligned} \quad (9.4)$$

$$Bt \leq |W|\mathbf{1} \quad (9.5)$$

$$\sum_{d \in D} s_d x_{l,d} \leq gt_l \quad l \in L \quad (9.6)$$

x binary variable and t integer variable.

We refer the above ILP problem as the GR ILP problem. We now present the following result:

Proposition 1 *If x_d and y_w are feasible solutions for the GRWA ILP problem, then x_d and t are feasible solutions for the GR ILP problem, where $t = \sum_{w \in W} y_w$.*

Proof. We first note that by summing over $w \in W$ in (9.2) it leads to (9.5). Secondly, (9.3) is the same as (9.6). Hence, the result follows. \square

Based on Proposition 1, we have

Proposition 2 *If x_d^* and t^* are the optimal solutions of the GR ILP problem, and there exists a binary y_w^* such that $\sum_{w \in W} y_w^* = t^*$ and $By_w^* \leq \mathbf{1}$ for $w \in W$, then x_d^* and y_w^* are the optimal solutions of the GRWA ILP problem.*

Proof. Suppose x_d and y_w are feasible solutions for the GRWA ILP problem, then based on Proposition 1, x_d and $t = \sum_{w \in W} y_w$ are feasible solutions for the GR ILP problem. Since x_d^* and t^* are the optimal solutions of the GR ILP problem, we have $\sum_{w \in W, l \in L} y_{l,w}^* = \sum_{l \in L} t_l^* \leq \sum_{l \in L} t_l = \sum_{w \in W, l \in L} y_{l,w}$. Therefore, the conclusion follows. \square

Obviously, the GR ILP problem is much easier to solve than the GRWA ILP problem since it has fewer integer variables and fewer constraints (e.g., see numerical examples in Section 9.4). More importantly, we can now relax the integer constraint on t in the GR ILP problem and solve a relaxed mixed ILP problem and then round up the values of t to obtain a solution for the GR problem. This would dramatically improve the computational efficiency. On the other hand, the relaxation approach is much less effective for the GRWA ILP problem since all its decision variables are binary. In general, if most lightpaths have relatively high wavelength counts (i.e., the values of their corresponding components in t are large), then the relaxed GR ILP problem often produces very good solutions for the GR problem, as illustrated by our numerical examples in Section 9.4. This is simply because if the optimal value of t_l is large, then the error of rounding up is relatively small.

9.3.2 The WA Problem

The WA problem of our interest is to find a binary solution y such that

$$\sum_{w \in W} y_w = t \quad \text{and} \quad By_w \leq \mathbf{1} \quad \text{for } w \in W,$$

where t is a feasible (or optimal) solution of the GR problem. This problem can be viewed as an ILP problem (without an objective function), which is much easier to solve than the GRWA ILP and the (relaxed) GR ILP problems. For example, it can be solved for networks with a few hundred nodes and lightpaths in seconds or minutes by using commercially available LP software, e.g., CPLEX. Based on Proposition 2, we know that if x and t are optimal solutions of the GR problem and the WA problem has a feasible solution y , then x and y are optimal solutions of the GRWA problem. In case when we cannot find a feasible solution for the WA problem, we can either increase the number of wavelengths in W in the WA problem (note that we can always find a feasible solution for the WA problem if W has enough wavelengths), or we can use $W^* \subset W$ in the GR problem (specifically, replace $|W|$ with $|W^*|$ in (9.5)) but still use W in the WA problem. Obviously, the latter approach is preferred in which case the decomposition method provides a feasible solution for the GRWA problem. An alternative approach is to use wavelength conversion via lightpath regeneration, which is equivalent to modifying L by breaking some lightpaths into two or more lightpaths. In addition, there are other possible remedies available to alleviate the infeasibility of the WA problem.

Though the WA problem can be solved as an ILP problem, it is also possible to solve it directly based on some heuristic algorithms (e.g., see [3]). In what follows, we consider a special type of the GRWA problem, in which the lightpaths satisfy a certain condition. Under such a condition, we show that a feasible solution for the corresponding WA problem can always be found, and we also develop an algorithm for finding a feasible solution. Without loss of generality, we assume that the capacity of every lightpath is one wavelength (i.e., $t_l = 1$ for every $l \in L$). For a lightpath whose capacity is more than one wavelength, we can treat it as several identical parallel lightpaths, each of which has capacity of one wavelength. Let p_e ($e \in E$) be the number of lightpaths that traverse fiber span e , and $p = \max_{e \in E} p_e$, which is the minimum number of wavelengths required for the network.

Define:

$$E_l := \{e \in E \mid e \text{ is on lightpath } l\}, l \in L;$$

$$L_e := \{l \in L \mid l \text{ traverses fiber span } e\}, e \in E.$$

We now present the following algorithm for the WA problem.

Algorithm 1 (for the WA problem)

1. Select an initial lightpath $l_0 \in L$ (arbitrarily), and assign a wavelength to l_0 .
2. Suppose $E_{l_0} = \{e_1, e_2, \dots, e_k\}$. Set $L_0 = \{l_0\}$. For $i = 1$ to k , do
 - a) Assign a wavelength to every lightpath $l \in L_{e_i} \setminus \cup_{0 \leq j < i} L_j$ such that no two lightpaths in L_{e_i} share the same wavelength (note that $L_{e_i} \setminus \cup_{0 \leq j < i} L_j$ is a subset of lightpaths in L_{e_i} to which wavelengths have not been assigned yet).
 - b) Let

$$L_i = L_{e_i} \setminus \cup_{0 \leq j < i} L_j,$$

$$E_i = \cup_{l \in L_i} E_l \setminus \{e_i\}.$$

We note that L_i is the set of lightpaths to which wavelengths are assigned in Step 2(a) and E_i is the set of fiber spans that are on at least one lightpath in L_i (excluding fiber span e_i).

3. For $i = 1, 2, \dots, k$, apply the procedure in Step 2 to E_i (with E_{l_0} being replaced with E_i), and continue until all the lightpaths in L are assigned (note that, since all the fiber spans in E_{l_0} have been considered already in Step 2, we can simply replace E_i by $E_i \setminus E_{l_0}$).

To study some useful properties associated with Algorithm 1, we first introduce the following terminologies:

Definition

1. We say a lightpath l and a fiber span e are connected (via fiber spans $\{e_1, \dots, e_m\}$ and lightpaths $\{l_1, \dots, l_m\}$) if there exist a set of fiber spans $\{e_1, \dots, e_m\}$ and a set of lightpaths $\{l_1, \dots, l_m\}$ such that $e_i \in E_{l_{i-1}}$ for $i = 1, \dots, m+1$ (where $l_0 \equiv l$ and $e_{m+1} \equiv e$) and $l_i \in L_{e_i}$ for $i = 1, \dots, m$.

2. We say two lightpaths l_0 and l_m are connected (via fiber spans $\{e_1, \dots, e_m\}$ and lightpaths $\{l_1, \dots, l_{m-1}\}$) if there exist a set of fiber spans $\{e_1, \dots, e_m\}$ and a sequence of lightpaths $\{l_1, \dots, l_{m-1}\}$ such that $e_i \in E_{l_{i-1}}$ and $l_i \in L_{e_i}$ for $i = 1, \dots, m$.
3. We say a set of lightpaths $\{l_1, \dots, l_m\}$ is a lightpath cycle if $E_{l_i} \cap E_{l_{i+1}} \neq \emptyset$ (i.e., lightpaths l_i and l_{i+1} share at least one common fiber span) for $i = 1, \dots, m$ ($l_{m+1} \equiv l_1$).
4. We say a lightpath cycle $\{l_1, \dots, l_m\}$ is a complete lightpath cycle if $E_{l_1} = \dots = E_{l_m}$, otherwise it is a non-complete lightpath cycle.

To help understand what is a lightpath cycle, consider the network depicted in Fig. 9.1. The network has four nodes (A, B, C, D), three fiber spans (A–B, B–C, B–D), and three lightpaths (A-B-C, C-B-D, D-B-A). It is clear that the three lightpaths (A-B-C, C-B-D, D-B-A) constitute a lightpath cycle, however it is a non-complete cycle.

We now present the following properties associated with Algorithm 1.

Proposition 3

1. Every fiber span in E_i is on at least one lightpath in L_i ;
2. For $1 \leq j \leq i$, $e_j \notin E_i$;
3. $L_i \cap L_j = \emptyset$ ($i \neq j$);
4. If $l \in L_i$, then it does not traverse fiber spans $\{e_1, \dots, e_{i-1}\}$;
5. If $E_i \cap E_j \neq \emptyset$ ($i \neq j$), then there exists a lightpath cycle with one lightpath in L_i and one lightpath in L_j ;
6. If a lightpath in L_i is connected to another lightpath in L_j in two different ways via lighpaths in $L \setminus \cup_{0 \leq h \leq k} L_h$ and fiber spans in $E \setminus E_{l_0}$, then there exists a lightpath cycle $\{l_1, \dots, l_m\}$ such that $E_i \cap (E_{l_{i_1}} \cap E_{l_{i_1+1}}) \neq \emptyset$ and $E_j \cap (E_{l_{i_2}} \cap E_{l_{i_2+1}}) \neq \emptyset$, where $1 \leq i_1 < i_2 \leq m$.

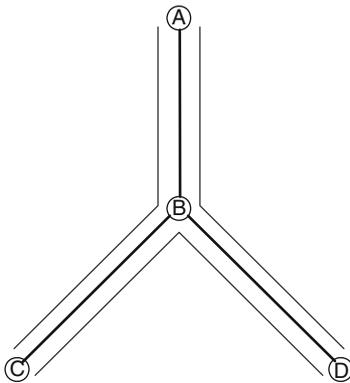


Fig. 9.1. A 4-Node network

Proof. We want to reiterate the fact that wavelengths are assigned to lightpaths in $L_i \subset L_{e_i}$ in Step 2(a).

1. By definition.
2. By definition, $e_i \notin E_i$. For $1 \leq j < i$ and $e \in E_i$, it is clear that we have assigned wavelengths to all the lightpaths in L_{e_j} before Step 2(a) while at least one lightpath in L_e has not been assigned by a wavelength before Step 2(a). Hence, $e_j \notin E_i$.
3. All the lightpaths in L_j are assigned by wavelengths at the end of Step 2(a) and they will not be considered again in later iterations.
4. By the same argument as in (2).
5. Suppose $e \in E_i \cap E_j$. Based on (1), e is on one lightpath in L_i , say l_i , and on another lightpath in L_j , say l_j . Furthermore, l_i and l_j traverse e_i and e_j , respectively, which are both on lightpath l_0 . Therefore, we have a lightpath cycle $\{l_i, l_0, l_j\}$.
6. The same argument used in (5) can be applied here as well. \square

In general, one needs to be careful about what wavelengths to use in Step 2(a) of Algorithm 1, otherwise it is possible that it may not produce a feasible solution for the WA problem. For example, consider the following example in which $E_{l_0} = \{e_1, e_2\}$, $L_1 = \{l_1\}$, $L_2 = \{l_2\}$, and $E_1 = E_2 = \{e\}$. If we assign the same wavelength to l_1 and l_2 , then we end up with assigning one wavelength to l_1 and l_2 on fiber span e , which is not permissible. Therefore, we have to assign u_1 and u_2 with different wavelengths.

It is clear that the number of different wavelengths needed in the WA problem is at least p . In what follows, we provide a sufficient condition under which p different wavelengths are enough to solve the WA problem.

Theorem 1 *If a network does not contain any non-complete lightpath cycle, then Algorithm 1 can produce a feasible solution for the WA problem which only needs p wavelengths.*

Proof. Since the network does not contain any non-complete lightpath cycle, based on (4), (5), and (6) in Proposition 3, we have (i) $E_i \cap E_j = \emptyset$ and (ii) no lightpath in L_i is connected to lightpath in L_j ($i \neq j$). Hence, when doing wavelength assignment for lightpaths in L_i in Step 2(a) we can use arbitrary wavelengths, and it guarantees that it is permissible (i.e., no two lightpaths that traverse the same fiber span would be assigned to the same wavelength). By repeating this argument, we can show that, in Algorithm 1, we can use arbitrary wavelengths in Step 2(a) and obtain a feasible solution for the WA problem. Since wavelengths used in Step 2(a) can be arbitrary, the maximum number of different wavelengths needed throughout Algorithm 1 should be no more than p . This completes our proof. \square

Theorem 1 implies that if a network does not contain any non-complete lightpath cycle, we can find a solution for the WA problem which only needs p wavelengths. In [3], the problem of whether the WA problem can be solved

with p wavelengths was also studied. However, we believe that the result there (Theorem 2 in [3]) is incorrect, which states that if a network is acyclic then its WA problem can be solved with p wavelengths. The network in Fig. 9.1 is a counter-example to this result. It is a tree (hence acyclic). Clearly we have $p = 2$, but need three wavelengths for its WA problem.

Since t in the WA problem is a feasible solution for the GR problem, i.e., $Bt \leq |W|\mathbf{1}$, we have $p \leq |W|$. This, together with Theorem 1, leads to the following result:

Theorem 2 *If a network does not contain any non-complete lightpath cycle, Algorithm 1 produces a feasible solution for the WA problem, and the decomposition method gives an optimal solution for the GRWA problem.*

To test whether a network contains any non-complete lightpath cycle, one can obviously use the exhaustive search method: finding all lightpath cycles and then test if any of them is non-complete. Clearly the complexity of this exhaustive search method grows exponentially. Currently, we do not have an efficient method to verify if a network contains any non-complete lightpath cycle. In fact, this problem itself could be NP-complete, just like the wavelength assignment problem.

In the case that a network contains non-complete lightpath cycles, let c^* be the minimum number of lightpaths that need to be removed from the network so that the remaining portion of the network does not contain any non-complete lightpath cycles. Then we have

Theorem 3 *There exists a feasible solution for the WA problem which requires at most $c^* + p$ wavelengths. Therefore, if $c^* + p \leq |W|$, then we can find a feasible solution for the WA problem and the decomposition method still gives an optimal solution for the GRWA problem.*

Before closing this section, we should point out that if the result in Theorem 3 can be further refined, then it can lead to better upper bounds on the number of wavelengths required for the WA problem.

9.4 Numerical Results

In this section, we present four sets of numerical examples. All ILPs and mixed ILPs were solved by using CPLEX 7.0 on a Dell Precision 420 PC with two 1GHz processors. We compare the numerical results obtained based on the three methods proposed in the previous two sections: the GRWA ILP formulation, the decomposition method combined with the GR ILP formulation, and the decomposition method combined with the relaxed GR ILP formulation. The run time for the decomposition method includes the run times for both the (relaxed) GR ILP problem and the WA problem. The run time for the WA problem in all four examples is very fast (it is less than a second in the

first three cases and less than 3 seconds in the last case). Our numerical results clearly indicate that the decomposition method combined with the relaxed GR ILP formulation produces quite good results with reasonably small run times.

Example 1. This is a relatively small network with 12 nodes, 17 fiber spans, 24 lightpaths, and 104 traffic demands (with different sizes). For this example, we were able to obtain the optimal solution based on the GRWA ILP formulation. The results are presented in Table 9.1.

Table 9.1. Numerical results for Example 1

	Run Time	Solution
GRWA ILP	400 seconds	128
GR ILP	80 seconds	128
Relaxed GR ILP	2 seconds	136

Example 2. The network we consider in this example has 30 nodes, 38 fiber spans, 47 lightpaths, and 242 demands (with different sizes). The results are presented in Table 9.2. For the GRWA ILP problem, we stopped the CPLEX program after 75 hours and obtained a feasible solution with objective value 249.

Table 9.2. Numerical results for Example 2

	Run Time	Solution
GRWA ILP	>75 hours	249
GR ILP	37 hours	189
Relaxed GR ILP	12 seconds	202

Example 3. The network in this example has 49 nodes, 75 fiber spans, 155 lightpaths, and 238 demands (with different sizes). It is a medium size network. For this example, the decomposition method based on the relaxed GR ILP problem produced a solution with value 345 in about 13 minutes, while the CPLEX program did not even return a feasible solution for the GRWA ILP and GR ILP problems after 40 hours (at which point we stopped the program). From the CPLEX program, we were also able to obtain a lower bound (based on the GR ILP problem) 328 for the objective function. Hence, the solution provided by the relaxed GR ILP-based decomposition method is within 5% of the lower bound. We note that the WA problem was solved in 0.37 seconds for this example. The results are presented in Table 9.3.

Table 9.3. Numerical results for Example 3

	Run Time	Solution
GRWA ILP	>40 hours	No Solution
GR ILP	>40 hours	No Solution
Relaxed GR ILP	13 minutes	345

Example 4. The network in this example has 144 nodes, 162 fiber spans, 299 lightpaths, and 600 demands (with different sizes). It is a relatively large network (a typical size for a nation-wide network). For this example, the method based on the relaxed GR ILP problem produced a solution in about 38 minutes, and the CPLEX program did not even return a feasible solution for the GRWA ILP and GR ILP problems after 100 hours (at which point we stopped the program). The WA problem in this case was solved in 2.67 seconds for this example. The results are shown in Table 9.4.

Table 9.4. Numerical results for Example 4

	Run Time	Solution
GRWA ILP	>100 hours	No Solution
GR ILP	>100 hours	No Solution
Relaxed GR ILP	38 minutes	431

References

1. A. Aggarwal, A. Bar-Noy, D. Coppersmith, R. Ramaswami, B. Schieber, and M. Sudan, "Efficient routing and scheduling algorithms for optical networks," In *Proceedings of the 5th ACM-SIAM Symposium on Discrete Algorithms*, pp. 412–423, Jan. 1994.
2. V. Auletta, I. Caragiannis, C. Kaklamanis, and P. Persiano, "Random path coloring on binary trees," In *Proceedings of the 3rd International Workshop on Approximation Algorithms for Combinatorial Optimization Problems (APPROX'00)*, LNCS 1913, Springer, pp. 60–71, 2000.
3. I. Chlamtac, A. Ganz, and G. Karmi, "Lightpath communications: an approach to high bandwidth optical WAN's," *IEEE Transactions on Communications*, Vol. 40, No. 7, pp. 1171–1182, 1992.
4. L. A. Cox, J. Sanchez, and L. Lu, "Cost savings from optimized packing and grooming of optical circuits: mesh versus ring comparisons," *Optical Networks Magazines*, pp. 72–90, May/June, 2001.
5. Z. Ding and M. Hamdi, "Clustering techniques for traffic grooming in optical WDM mesh networks," *Proceedings of 2002 Global Telecommunications Conference*, Vol. 3, pp. 2711–2715, 2002.

6. W. Grover, J. Doucette, M. Clouqueur, D. Leung, and D. Stamatelakis, "New options and insights for survivable transport networks," *IEEE Communications Magazine*, Vol. 40, No. 1, pp. 34–41, January 2002.
7. J. Q. Hu and B. Leida, "Traffic Grooming, Routing, and Wavelength Assignment in Optical WDM Mesh Networks," *Proceedings of 2004 INFOCOM*, pp. 495–501, 2004.
8. C. Lee and E. K. Park, "A genetic algorithm for traffic grooming in all-optical mesh networks," *Proceedings of 2002 IEEE International Conference on Systems, Man and Cybernetics*, Vol. 7, 2003.
9. M. Mihail, C. Kaklamanis, and S. Rao, "Efficient access to optical bandwidth–wavelength routing on directed fiber trees, rings, and trees of rings," In *Proceedings of the 36th Annual Symposium on Foundations of Computer Science*, pp. 548–557, Milwaukee, Wisconsin, 23–25 October, 1995.
10. K. G. Patrocínio Jr. and G. R. Mateus, "A lagrangian-based heuristic for traffic grooming in WDM optical networks," *Proceedings of 2003 Global Telecommunications Conference*, Vol. 5, pp. 2767–2771, 2003.
11. P. Raghavan and E. Upfal, "Efficient routing in all-optical networks," In *Proceedings of the 26th ACM Symposium on Theory of Computing*, pp. 134–143, May 1994.
12. R. Ramaswami and K. Sivaraman, *Optical Networks: A Practical Perspective*, Morgan Kaufmann Publishers, 1998.
13. S. Subramaniam, M. Azizoglu, and A. Somani, "On optimal converter placement in wavelength-routed networks," *IEEE/ACM Transactions on Networking*, Vol. 7, No. 5, pp. 754–766, 1999.
14. Y. Tao, *Design and Performance Analysis of WDM Optical Networks*, Ph.D. Dissertation, Boston University, 2006.
15. Y. Tao and J.Q. Hu, "Amplifier Placement for Optical Networks," Manuscript, Department of Manufacturing Engineering, Boston University, 2003.
16. K. Zhu and B. Mukherjee, "Traffic grooming in an optical WDM mesh network," *Proceedings of IEEE International Conference on Communications*, pp. 721–725, Helsinki, Finland, June, 2001.
17. "The intelligent optical core: improving bandwidth efficiency," Sycamore Networks, white paper, 2001.

Survivable Traffic Grooming

Arun K. Somani

10.1 Introduction

Optical fiber medium using wavelength division multiplexing (WDM) offers tremendous transmission bandwidth to deliver high-bandwidth services cost effectively. A WDM-based network divides the fiber capacity into non-overlapping wavelength channels, each of which operates at a transmission rate compatible with electronics. The routing function is controlled by the optical layer management planes. The nodes equipped with *optical cross-connects* support routing by switching different wavelengths at input ports to different output ports. The switching is managed by the control plane. An end-to-end path, called a *lightpath*, is established using the same wavelength on all links on the chosen path, or a different wavelength on a different link on the path with required *wavelength conversion* provided by intermediate links on the path. An all optical domain on path can provide complete transparency in data transmission. A network layer, called the WDM layer, in a layered architecture provides interfaces to other layers such as SONET/SDH gear, IP (Internet Protocol) ATM (Asynchronous Transfer Mode) or any other transport technology to connect to a lightpath.

10.1.1 Protection and Restoration

As wavelength routing paves the way for network throughput of possibly hundreds of Tb/s, network survivability assumes critical importance. A loss or damage to a fiber is a common means of a greater loss. A short network outage can lead to huge data loss. Thus a connection being carried in the network also needs high protection. Survivability refers to the ability of the network to reconfigure and reestablish communication upon failures. According to industry standards, the expected availability requirements are 99.99% or higher. The basic types of network failures generally considered are link and node failure. Cable cuts that cause link failures are common in optical networks. A node failure is usually due to equipment failure at the node. Channel failure

is unique to WDM networks. A channel failure is usually caused by the failure of transmitting or receiving equipment operating on that channel.

Several survivability techniques have been proposed [1, 2, 3] and can be classified into two general categories: preplanned protection and dynamic restoration. In preplanned protection-based techniques resources are already planned, typically at the time of establishing a connection, to recover from network failures and hence recovery is faster. During the operation phase these reserved resources remain idle. Upon the occurrence of failure, reserved resources are used to recover from the failure according to protection protocols. In contrast, in dynamic restoration, the resources used for recovery from failure are not reserved at the time of connection establishment, but are discovered dynamically when a failure occurs. As is obvious, dynamic restoration uses resources efficiently, but the restoration time is usually longer. Moreover, 100% service recovery cannot be guaranteed as it is not guaranteed that the spare capacity is available at the time of failure. To guarantee the service, a preplanned protection is the preferred approach. Most researchers therefore focus on preplanned protection. Figure 10.1 shows a classification of survivable design techniques that can be deployed at various layers.

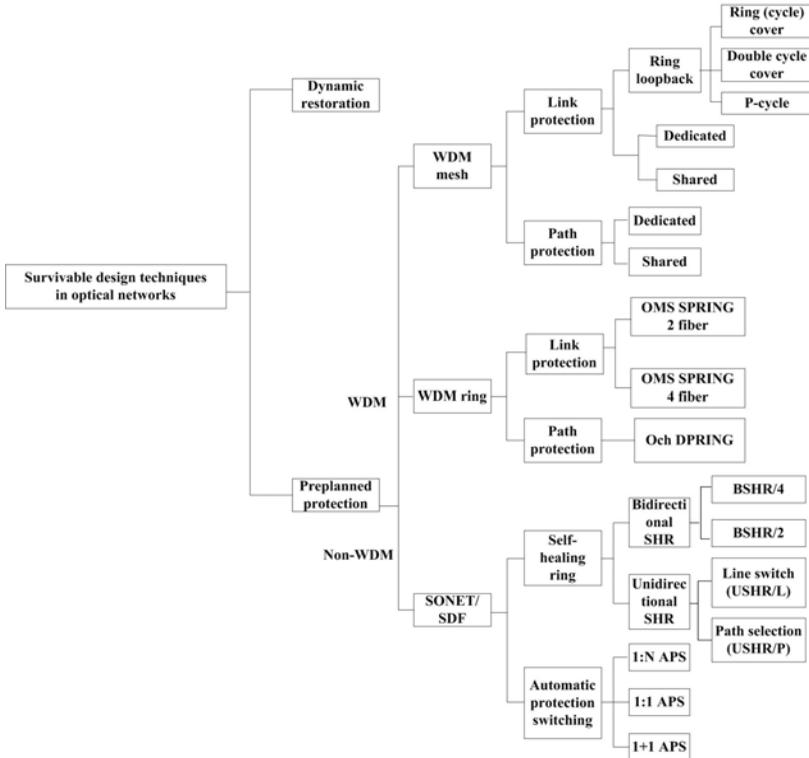


Fig. 10.1. Survivable design techniques in optical networks

Since the protection can be provided at many layers, a relevant question is, “which layer is the best?”. Although there is no consensus, most researchers believe that the optical layer protection is the best. It is very efficient in handling failures such as fiber cuts, which are probably the biggest cause of failure. To provide protection at any other layer requires that the layer must keep track of all logical links that may be affected by a single fiber failure. Notice that a logical path may use many lightpaths and they may be routed on a fiber/link. This implies that the relationship between established routes and fiber links must be collected and managed. This imposes additional overheads on protocols, algorithms, and processing during the normal operation. It is also possible that such an information cannot be made available by the other service provider. The survivability of managing layer, upon detection of a failure, may need to establish additional logical links upon failure that may lead to delay. Moreover, preplanning may require establishing logical links and/or lightpaths that may or may not use certain fiber links. Thus, in planning for a failure, the management, and actual recovery may overwhelm the management system. Moreover, optical layer protection and restoration provides an additional resilience in the network. For example, optical layer can be used to provide resilience against multiple failures. Furthermore, the layers above the optical layer may not be fully able to provide the protection. If protection is provided at many layers, then significant cost savings can be obtained by making use of optical layer protection and restoration.

It should, however, be clear that optical layer protection cannot handle any faults in the higher layer of network. Because of the property of protocol transparency, the optical layer may be unaware of what exactly is carried on the lightpaths, and therefore, cannot monitor the traffic to sense any degradation. A protection mechanism must also consider multiple, in particular two, link failure scenarios due to common routing of fiber cables.

10.2 WDM Mesh Network Protection

Survivable design in optical mesh networks involve high efficiency in capacity utilization and fast restoration. The problem is complex because there are multiple routes that can be used to recover from failures. A path that is used to carry the traffic under normal conditions is called a *primary* or a *working* path. A path that is reserved for protection is called a *backup* or a *protection* path. The problem becomes more complex when each wavelength is also groomed with smaller traffic streams as the number of affected traffic streams due to a single link failure is much larger than a wavelength-based connection. Protection schemes in mesh networks are broadly classified as either link-based or path-based, as shown in Fig. 10.2.

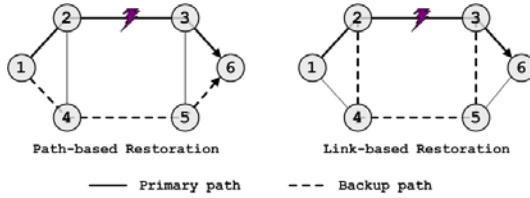


Fig. 10.2. Link- and path-based protection

A path-based method employs end-to-end detouring. The traffic is rerouted by the source node of affected path to the destination. A pair of link-disjoint paths are established between the source and destination nodes of each connection request. One path is used as the working path, and the other is used as the backup path. When a link failure occurs, the source node of the affected path switches the traffic to its backup path. In link-based protection, backup paths for each link are pre-computed. Upon the failure of a link, the working connections on the link are switched by the two end nodes of the link to their corresponding backup paths. Thus, a link-based method employs local detouring and the traffic is rerouted around the failed link.

An alternate link based protection scheme is called a subgraph-based routing [4]. The idea behind the *subgraph-based routing* scheme is to plan network resource utilization in such a way that, for any link failure, there exists an alternate path for every accepted request. When a link fails, all paths that get affected by the link failure are reassigned to their new paths.

The link- and path-based protection schemes can use either dedicated or shared resources. Dedicated resources are exclusively reserved for backup paths. In case of dedicated protection, the optical cross-connects (OXC) for the backup path can be preconfigured requiring only the source node to switch the path. If two or more primary paths are link disjoint, then the corresponding backup paths can share resources as only one of the primary paths will fail due to a single link failure. This is called backup multiplexing. This utilizes fewer resources, but requires a more complicated management. In shared backup path-based protection (SBPP), it is necessary to configure OXC for the backup path accordingly after a link failure occurs. This inevitably increases the recovery delay. SBPP tends to use lesser total capacity than shared link-based protection.

The link-based methods limits the choices for alternatives. However, link-based protection tends to be faster than shared path-based protection because only the two end nodes of the failed link are involved in recovery. One drawback for link-path protection in WDM networks, however, is that the backup path must necessarily use the same wavelength as the primary path. To tolerate node failures, a link based scheme can be extended to a two-link segment-based protection scheme so that the path can be routed around a failed node.

10.2.1 Alternate Link-Based Protection Schemes

A few special types of link-based protection schemes have also been developed, which aim to benefit by using the fast restoration of ring-like protection or save resources. Notable among these are: *Double Cycle Cover*, the *p-Cycle*, and subgraph-based routing schemes.

In the double cycle cover [5], the network is represented by a directed graph. A set of cycles are embedded on the given topology. Each link is covered by two directed cycles, one in each direction. For planar graphs, the required set of protection cycles can be found in polynomial time, but no known polynomial-time algorithm for non-planar graphs is known [5]. The model uses fiber-based recovery, similar to the SONET BLSR ring. On each link, half of the capacity is reserved for backup and the other half is used for working traffic in each direction to protect the traffic in opposite direction. Since the protection switches can be preconfigured, this method can achieve fast restoration, but at the cost of 100% redundancy.

The *p-Cycle* protection method [6, 7] also uses a cyclic layout of spare capacity to provide protection. When a link fails, only the nodes neighboring the failure need to perform real-time switching. The key difference between the *p-Cycle* protection and the ring cycle protection, such as double cycle cover, is that the *p-Cycle* protection not only protects the links on the cycle, as in the ring protection, it also protects straddling links. A straddling link is an off-cycle link whose two end nodes are both on the cycle. This important property effectively improves the capacity efficiency of *p-Cycles*.

Figure 10.3 depicts an example that illustrates *p-Cycle* protection. In Fig. 10.3 (a), A-B-C-D-E-A is a *p-Cycle* formed using spare capacity. When an on-cycle link A-B fails, the *p-Cycle* provides protection as shown in Figure 10.3(b) for traffic from A to B by using the path A-E-D-C-B. A reverse path is followed for traffic in the other direction, i.e., path B-C-D-E-A is used. When a straddling link B-D fails, the *p-Cycle* provides two backup paths, namely B-A-E-D and B-C-D for traffic from B to D as shown in Fig. 10.3(c). Thus the full capacity of a straddling link can carry working traffic, as each backup path is able to provide half of the capacity. For traffic in the other direction, i.e., D to B, paths are used in the reverse direction. Thus straddling links are covered fully by a cycle whereas an on-cycle link can use only half the capacity for primary connections, the other half has to be reserved for protection.

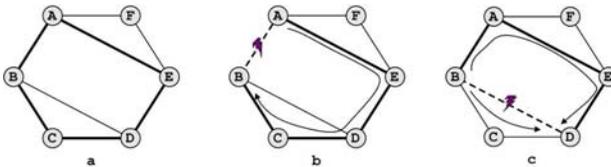


Fig. 10.3. (a) A *p-Cycle*. (b) protecting on-cycle link. (c) protecting straddling link

The subgraph-based routing strategy [4] provides a passive form of redundancy in the event of a fault occurrence from a given set of faults. The routing algorithm maintains $|F| + 1$ copies of states of the network resource utilization where $|F|$ is the number of possible faults. The i th copy of the state, S_i , represents the state of network when all accepted requests are routed without using the failed resources due to the presence of fault f_i . One copy of the network state, S_0 , represents a fault free state. When a request arrives, it is accepted only if it can be routed in each of $|F| + 1$ states using the free resources in the corresponding state. If a request is accepted, the resources used in each state are marked as busy in that state. If a fault f_i occurs, the state S_i provides the resources that must be used to route all accepted requests.

The scheme is passive in the sense that resources are available, but not reserved or preconfigured at the time of setting up connections. An end user experiences nominal interruption in service when a fault occurs and the connections are rerouted/restored. A system process to carry out restoration is described in [8]. The key characteristics of this scheme are that no additional resources are used to provide connection redundancy, although a 100% guarantee against all known faults is provided using a path-based fault tolerance strategy for anticipated failures. It is a path based recovery strategy because it does not guarantee that any of the same links are used to reroute a connection upon the occurrence of a fault. A disadvantage is that the scheme can potentially require a complete reconfiguration of the network to a predetermined new state. However, it is shown to be most resource efficient [4].

10.2.2 Dynamic Traffic Handling

Dynamically provisioning connections on demand is becoming more important in groomed transport networks. Early research has mostly focused on static traffic. Dynamic traffic can be handled by either using a strategy where the primary and backup paths are determined when a request arrives or backup resources are reserved and a connection is accepted only if both primary and backup paths can be found among the respective available resources for each. The objective of partitioning the resources is to guarantee that the capacity available for routing randomly arriving connection requests will be 100% protected by the reserved protection capacity. A p -Cycle based scheme can also deal with dynamic traffic using the latter scheme of resource reservation and a two-step approach. The basic idea is to provision the network resources in two parts: protection resources and working-capacity resources in the p -Cycles. The design must also ensure that the p -Cycles are preconfigured. In the first step, one can find a set of p -Cycles to cover the network and reserve enough capacity in p -Cycles. In the second step, one routes the requests as they arrive. The reserved capacity on p -Cycles leads to less control signaling overhead and less dynamic state information needs to be maintained to achieve protection. Thus, the p -Cycle based design has the advantage of fast recovery, less control signaling, and less dynamic state information to be maintained.

10.3 Protection Design: Optimization and Performance

The problem of restorable network design for a static traffic demand has been dealt with in [1, 2, 3, 9, 10, 11] and for dynamic traffic scenario in [12, 13, 14, 15, 16]. For protection mechanism design and performance analysis, a network is represented using a directed graph with N nodes and L links. For dedicated and shared path protection design, a set of K link-disjoint alternate routes for every source–destination pairs can be pre-computed. It has been shown that $K = 2$ or 3 is sufficient to achieve good performance to tolerate a single link failure. For a network with W wavelengths and no wavelength conversion, K link-disjoint alternate routes can be viewed as $K \times W$ paths, each of which is wavelength continuous path. In shared path protection scheme, the capacity optimization problem is to choose a pair of link-disjoint paths to serve as primary and backup for each connection request. The objective function is to minimize the total capacity required. The shared backup path protection (denoted as SBPP) can be implemented in two forms. In the first form, the goal is to minimize the total number of wavelengths used on all links when paths are not pre-configured. This is generalized to obtain the second scheme where the paths are pre-configured (denoted as PRE-SBPP).

An alternate approach is to use p -Cycle based protection (denoted as PCP), where the goal is to identify a set of simple distinct cycles that are sufficient to protect against all failures of interest. Longer cycles waste more capacity for redundancy and create long backup paths. A two-step algorithmic approach for the p -Cycle optimization, therefore, identifies a set of candidate p -Cycles in the first step and iteratively chooses a subset to deploy in actual implementations. Usually, the preselected set in the first step only includes cycles of up to a certain length. In a joint optimization design, one attempts to optimize the choice of routing working connections in conjunction with the p -Cycle selection, so that the total capacity is minimized.

In most cases, an Integer Linear Programming (ILP) formulation is used to choose optimal paths and/or cycles. We give an outline of ILP formulations for the three schemes, SBPP, PRE-SBPP, and PCP.

10.3.1 ILP Formulation

To formulate the ILP for the three cases, we use the following notations.

- $n = 1, 2, \dots, N$: Number assigned to each node
- $l, k = 1, 2, \dots, L$: Number assigned to each link
- $\lambda = 1, 2, \dots, W$: Number assigned to each wavelength
- $i, j = 1, 2, \dots, N(N - 1)$: Number assigned to s - d pair
- $K = 2$: Number of alternate routes between each s - d pair
- $p, r = 1, 2, \dots, KW$: Number assigned to a path
- (i, p) : Refers to the p th path for s - d pair i
- d_i : Demand for node pair i , in terms of number of lightpath requests
- ρ_l^n : It equals one if node n is an end node of link l , else zero (data)

The following notations are used for path information

Variable: $\delta^{i,p} = 1$ if (i, p) is chosen as a primary path, else 0

Variable: $\nu^{j,r} = 1$ if (j, r) is chosen as a restoration path, else 0

Data: $\epsilon_l^{i,p} = 1$ if link l is used in path (i, p) , else 0

Data: $\psi_\lambda^{i,p} = 1$ if λ is used by the path (i, p) , else 0

Variable: $g_{l,\lambda} = 1$ if λ is used by a restoration route that traverses link l

Variable: s_l : # of wavelengths used by backup lightpaths on link l

Variable: w_l : # of wavelengths used by primary lightpaths on link l

Data: $I_{(i,p)(j,r)} = 1$ if paths (i, p) and (j, r) share link. If $i = j$, then $p \neq r$

Data: $\Pi_{(i,p)(j,r)}^\lambda$: # of shared links between paths (i, p) and (j, r)

The following additional notations are used in formulation for PCP.

$c = 1, 2, \dots, P$: Number assigned to a cycle.

Data: $\omega_c^l = 1$ if link l is on cycle c ; else 0

Data: $\sigma_c^l = 1$ if link l is a straddling link on cycle c , else 0

Variable: $\tau_c^\lambda = 1$ if cycle c is chosen and uses wavelength λ , else 0

ILP Formulation for General Shared Path Protection

$$\text{OBJECTIVE :} \quad \text{Min} \sum_{l=1}^L (w_l + s_l) \quad (10.1)$$

1. *Link capacity constraint:*

$$w_l + s_l \leq W \quad 1 \leq l \leq L \quad (10.2)$$

2. *Demand constraint for each node pair:*

$$\sum_{p=1}^{KW} \delta^{i,p} = d_i \quad 1 \leq i \leq N(N-1) \quad (10.3)$$

3. *Primary link capacity constraint:* # of primary lightpaths on a link.

$$w_l = \sum_{i=1}^{N(N-1)} \sum_{p=1}^{KW} \delta^{i,p} \epsilon_l^{i,p} \quad 1 \leq l \leq L \quad (10.4)$$

4. *Spare capacity constraint:* Spare capacity required on link l .

$$s_l = \sum_{\lambda=1}^W g_{l,\lambda} \quad 1 \leq l \leq L \quad (10.5)$$

5. *Primary path wavelength usage constraint:*

$$\sum_{i=1}^{N(N-1)} \sum_{p=1}^{KW} \delta^{i,p} \epsilon_l^{i,p} \psi_\lambda^{i,p} + g_{l,\lambda} \leq 1 \quad 1 \leq l \leq L, 1 \leq \lambda \leq W \quad (10.6)$$

6. *Restoration path wavelength usage constraint:*

$$g_{l,\lambda} \leq \sum_{i=1}^{N(N-1)} \sum_{r=1}^{KW} \nu^{i,r} \epsilon_l^{i,r} \psi_\lambda^{i,r} \quad 1 \leq l \leq L, 1 \leq \lambda \leq W \quad (10.7)$$

$$N(N-1)KW g_{l,\lambda} \geq \sum_{i=1}^{N(N-1)} \sum_{r=1}^{KW} \nu^{i,r} \epsilon_l^{i,r} \psi_\lambda^{i,r} \quad 1 \leq l \leq L, 1 \leq \lambda \leq W \quad (10.8)$$

7. *Demand constraint for node pair i :* There is one restoration route for each primary call. Let $u \in \{0, 1\}$ and $v = 1 - u$:

$$\sum_{p=uW+1}^{(u+1)W} \delta^{i,p} = \sum_{r=vW+1}^{(v+1)W} \nu^{i,r} \quad 1 \leq i \leq N(N-1) \quad (10.9)$$

8. *Constraint for topology diversity of primary and backup paths:* Primary and backup paths should be link disjoint. Let $m, n \in \{0, 1\}$; $s = 1 - m$; and $t = 1 - n$. The primary path of a node pair can be any one of the two alternate paths for this node pair. We use m th path of node pair i as primary path for node pair i , n th path of node pair j as primary path of node pair j , s th path of node pair i as backup path of node pair i , and t th path of node pair j as backup path for node pair j . If $I_{(i,m)(j,n)} = 1$,

$$(\nu^{i,sW+\lambda} \psi_\lambda^{i,sW+\lambda} \epsilon_l^{i,s} + \nu^{j,tW+\lambda} \psi_\lambda^{j,tW+\lambda} \epsilon_l^{j,t}) I_{(i,m)(j,n)} \leq 1 \quad (10.10)$$

$$1 \leq i, j \leq N(N-1), 1 \leq \lambda \leq W, 1 \leq l \leq L$$

Pre-Cross-Connected Shared Path Protection

All other constraints for this formulation remain the same except that the pre-cross-connected protection requires an additional constraint. Suppose link l , k , and m share a common node n . If backup path b_1 uses link l and k , and backup path b_2 uses link l and m , then b_1 and b_2 cannot share the backup capacity on link l .

$$(\nu^{i,p} \epsilon_l^{i,p} \epsilon_k^{i,p} \psi_\lambda^{i,p} + \nu^{j,r} \epsilon_l^{j,r} \epsilon_m^{j,r} \psi_\lambda^{j,r}) \rho_l^n \rho_k^n \rho_m^n \leq 1 \quad (10.11)$$

$$1 \leq i, j \leq N(N-1), 1 \leq p, r \leq KW, 1 \leq l, k, m, \leq L, 1 \leq n \leq N$$

P-Cycle Protection

The objectives, link capacity constraint, demand constraint, and primary link capacity constraint remain as above. The other constraints are as follows.

Spare capacity constraint:

$$s_l = \sum_{c=1}^P \sum_{\lambda=1}^W \omega_c^l \tau_c^\lambda \quad 1 \leq l \leq L \tag{10.12}$$

Primary path wavelength usage constraint: Only one primary path can use a wavelength λ on link l ; no p -Cycle can use the same λ on link l .

$$\sum_{i=1}^{N(N-1)} \sum_{p=1}^{KW} \delta^{i,p} \epsilon_l^{i,p} \psi_\lambda^{i,p} + \sum_{c=1}^P \omega_c^l \tau_c^\lambda \leq 1 \tag{10.13}$$

$$1 \leq l \leq L, 1 \leq \lambda \leq W \tag{10.14}$$

Restoration guarantee constraint for link l : There are enough p -Cycles and wavelengths to recover the failure of link l .

$$\sum_{i=1}^{N(N-1)} \sum_{p=1}^{KW} \delta^{i,p} \epsilon_l^{i,p} \psi_\lambda^{i,p} \leq \sum_{c=1}^P \sigma_c^l \tau_c^\lambda \tag{10.15}$$

$$1 \leq l \leq L, 1 \leq \lambda \leq W \tag{10.16}$$

Performance of Different Schemes

The capacity performance of the three protection schemes has been studied by experimenting on six topologies [15]. Figure 10.4 on left shows a Pan-European COST239 network with 11 nodes and 26 links with an average node degree of 4.7 when all links are present. To study the effect of average nodal degree of a network on the performance of different schemes, three more topologies are created by manipulating the original COST239 network by deleting 5 (dotted), additional 4 (dashed), and additional 3 (dot-dashed) edges sequentially. These edges are progressively minimally used for a traffic matrix shown in Fig. 10.5 on the left. This results into four topologies. Two more topologies used in the experiments are a 11-node 22-link NJ-LATA network, which has average nodal degree 4, and a 14-node 21-link NSFNET, which has an average nodal degree 3, as shown in Fig. 10.4 in middle and right, respectively.

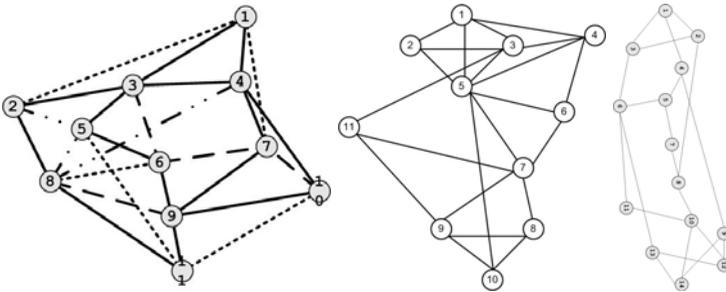


Fig. 10.4. 11N/26L COST 239, 11N/22L NJ LATA and 14N/21L NSFNET.

Node	1	2	3	4	5	6	7	8	9	10	11
1	0	1	1	3	1	1	1	1	1	1	1
2	1	0	5	8	4	1	1	10	3	2	3
3	1	5	0	8	4	1	1	5	3	1	2
4	3	8	8	0	6	2	2	11	11	9	9
5	1	4	4	6	0	1	1	6	6	1	2
6	1	1	1	2	1	0	1	1	1	1	1
7	1	1	1	2	1	1	0	1	1	1	1
8	1	10	5	11	6	1	1	0	6	2	5
9	1	3	3	11	6	1	1	6	0	3	6
10	1	2	1	9	1	1	1	2	3	0	3
11	1	3	2	9	2	1	1	5	6	3	0

Topology	Avr.		PRE-	
COST239	degee	SBPP	SBPP	PCP
26 links	4.7	816	948	794
Opt. Gap		$\leq 0.4\%$	$\leq 0.4\%$	$\leq 5.0\%$
21 links	3.8	899	1039	908
Opt. Gap		$\leq 3.0\%$	$\leq 2.0\%$	$\leq 3.0\%$
17 links	3.1	982	1108	1118
Opt. Gap		$\leq 3.0\%$	$\leq 2.0\%$	$\leq 0.4\%$
14 links	2.5	1280	1427	1717
Opt. Gap		$\leq 0.4\%$	$\leq 3.0\%$	$\leq 5.0\%$

Fig. 10.5. Traffic Matrix and Capacity used in terms of # of wavelength links

For the traffic matrix, the total capacity used by three different protection schemes in the four variations of COST 239 network are presented in Figure 10.5 on the right by solving the ILP formulation. As the average nodal degree increases gradually in the modified COST239 network from 14 links to 26 links, the total capacity required for establishing restorable connections for all the requests decreases for all three protection schemes. The decline of total capacity used is due to the decline of both working and backup capacity as the network connectivity increases. The improvement in the capacity utilization is because the alternate paths become shorter and the opportunity for backup capacity sharing in SBPP and PRE-SBPP increases as the network connectivity increases. The improvement of capacity efficiency with the increase in network connectivity is more dramatic in PCP scheme than in path-based protection (SBPP and PRE-SBPP), which is good for low-connectivity networks. The performance optimization gap in solving the ILP is also shown.

To further validate the reasoning and to study the effect of network connectivity on the capacity performance of three different protection schemes, two randomly generated traffic matrices with different characteristics are used to conduct experiments on the six network topologies. The number of requests in Type I traffic is smaller than that in Type II traffic. The average total capacity used for each protection scheme is the average value of ten sets of requests. Table 10.1 shows the average total used capacity by three different schemes in six topologies for Type I and Type II traffic, respectively.

The results indicate that general shared-path protection and pre-cross-connected shared-path protection use less total capacity than p -Cycle protection in low-connectivity networks, while they are comparable in high-connectivity networks. The p -Cycle protection scheme appears to be a better option for protection in high-connectivity networks. Pre-cross-connected shared-path protection uses more total capacity than general shared backup path protection. Pre-cross-connected shared-path protection can achieve faster restoration while remaining to be a path-based method.

Table 10.1. Average total capacity used by different protection schemes in six topologies for Type I and Type II traffic (number of wavelength links) for various topologies

Topology	Avr. degree	Type I Traffic			Type II Traffic		
		SBPP	PRE-SBPP	PCP	SBPP	PRE-SBPP	PCP
Cost239/26 links	4.7	72	74	70	101	106	93
NJ-LATA	4.0	78	81	83	111	118	114
Cost239/21 links	3.8	77	80	83	110	117	112
Cost239/17 links	3.1	91	97	110	130	142	155
NSFNET/21 link	3.0	105	113	132	147	162	181
Cost239/14 links	2.5	106	115	136	147	164	186

10.4 Traffic Grooming and Survivable Design

As noted in earlier chapters, grooming is the ability to share resources among multiple requests needing the network resource. This sharing is possible and even desirable because the resources under question are expensive, individual entities need only a fraction of the resource, and multiplexing allows the resource cost to be amortized over the number of users. Given the generic nature of the problem, grooming can be provided within a layer or across layers, such as optical grooming and electronic grooming. The optical layer is typically equipped with reconfigurable switching fabrics and the optical grooming technique used depends on the granularity and time scale of the switching functionality available in the optical layer. Optical switching can be done at various levels of granularity and time scales that include: waveband switching, wavelength switching, sub-wavelength time slot level switching, burst switching, flow switching on lightpaths, and packet switching.

Wavelength level switching techniques allow circuit switched sharing of a wavelength. When supplemented with some additional hardware an overlaid control protocol can allow statistical sharing of a wavelength as well. For example, in sub-wavelength time slot level switching, the wavelength is divided into T fixed time slots and the optical switching fabric is reconfigured for every k time slots ($k = 1..T$). Optical circuits are configured in the form of a path or a tree, and are categorized into four main classes:

1. Point to Point (P2P): this technique allows a lightwave circuit to aggregate traffic between the convenor node and the end node of the circuit.
2. Point to Multi-point (P2MP): this technique allows a lightwave circuit to aggregate traffic from a source to multiple destinations.
3. Multipoint to Point (MP2P): this technique allows a lightwave circuit to aggregate traffic from multiple sources to the same destination.
4. Multipoint to Multi-point (MP2MP): this technique allows a lightwave circuit to aggregate traffic from multiple sources to multiple destinations.

In all cases, the protection mechanisms are designed using the same general principles described earlier. The different traffic streams using the same primary path, however, may use very different backup paths in case of grooming. Also, a wavelength channel may carry both primary and backup capacity. These factors affect the overall utilization of resources and performance.

The study in [13] addressed the problem of dynamically establishing dependable low-rate traffic stream connections in WDM mesh networks with traffic grooming capabilities. To establish a dependable connection, they pre-compute link-disjoint primary and backup paths between the source and destination node and use backup multiplexing to reduce the overhead of backup traffic streams. Two schemes for grooming traffic streams onto wavelengths were proposed, namely *Mixed Primary-Backup Grooming Policy* (MGP) and *Segregated Primary-Backup Grooming Policy* (SGP). In SGP scheme, a wavelength is either used for primary paths or backup paths, but not both. MGP scheme does not put that constraint. The simulation results showed that SGP performs better in mesh networks and MGP performs better in ring networks.

In [17], the authors proposed three approaches, namely protection-at-lightpath (PAL) level, mixed protection-at-connection (MPAC) level, and separate protection-at-connection (SPAC) level, for grooming a connection request with shared protection. The authors concluded that when the lower bandwidth connections outnumber higher bandwidth connections, it is beneficial to groom working paths and backup paths separately, especially when the number of grooming ports is sufficient. Otherwise, protecting each specific lightpath achieves the best performance.

In the following, we investigate the problem of how to *groom* subwavelength level requests efficiently in mesh restorable WDM networks, and formulate the corresponding path selection and wavelength assignment problem as ILP optimization problems. Since the ILPs become too complex too soon, based on the same design principles, we also develop a heuristic algorithm for routing dynamic traffic in grooming networks. In some cases the network resources may not be sufficient to protect all traffic streams to the full extent to provide 100% protection. For such cases, our approaches of survivable grooming network design is also extended to partially protected WDM grooming networks.

10.4.1 Formulation of the Optimization Problem

We again use the same network model as in the previous section and with W wavelengths and $K = 2$ disjoint alternate paths for each s-d pairs, we have $W \times K$ networks, one each of a wavelength. The first W networks, numbered from 1 to W , contain the first alternate path for each s-d pair and the second W networks, numbered from $W + 1$ to $2W$, contain the second alternate path. Figure 10.6 (left) illustrates this layered model of a 6-node network with 3 wavelengths, 2 connections with each having 2 link disjoint alternate paths. Please note that if a path among network 1 to W is selected as a primary path for a request, its backup paths can only be selected from the network

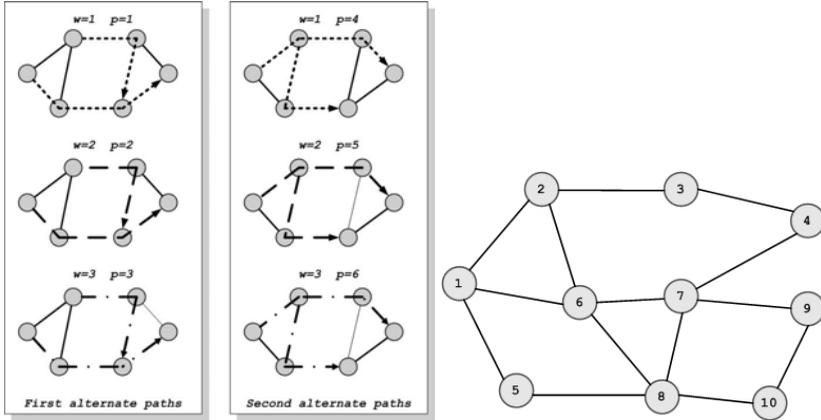


Fig. 10.6. (a) An example of layered network model with $W = 3$ and $K = 2$ and (b) an experimental topology

$W + 1$ to $2W$ in this layered network model. We consider 100% restoration guarantee for any single link failure for protected connections.

10.4.2 Shared Backup Reservation

In grooming WDM networks, the capacity reserved for restoration paths is more complicated. Let $B = \{b_1, b_2, \dots, b_k\}$ denotes the set of backup paths that traverse the wavelength w on link l . Let their respective capacities be $D = \{d_1, d_2, \dots, d_k\}$ and their respective primary paths be $P = \{p_1, p_2, \dots, p_k\}$. If none of the p_i 's have common links, the needed capacity on w is $\max(d_1, d_2, \dots, d_k)$. If some of the p_i 's have common links, their backup paths can still be groomed on wavelength w . However, the capacity to be reserved must be up to the summation of their capacities. The primary paths can be grouped according to their common links. Let $P^l = \{p_1^l, p_2^l, \dots, p_a^l\}$ denote the group of primary paths that have link l as their common link. The capacity required by this group for back up of link l is then given by $D^l = (d_1^l + d_2^l + \dots + d_a^l)$. It is possible that one primary path belongs to more than one group. The reserved capacity on wavelength w on link l is therefore the maximum value of the capacities required by all the groups, that is $D = \max(D^l)$.

10.4.3 Dedicated Backup Reservation

One simple and effective way of assigning backup capacities is to reserve dedicated capacity for each backup path. While choosing primary paths, instead of simply choosing the shortest path, we try to *minimize* the total *link-primary-sharing* (MLPS). The link-primary-sharing is defined as $s_l = \max(0, P_l - 1)$

where s_l denotes the link-primary-sharing of link l and P_l denotes the total number of primary paths that utilize link l . s_l can be viewed as the penalty assigned to link l when it is used by more than one primary path.

We formulate backup multiplexing and dedicated backup reservation schemes with MLPS goals as ILP optimization problems below.

To formulate the grooming survivable network design problem in a WDM mesh network with static traffic patterns as an ILP problem, we assume that the network is a single-fiber general mesh network. A connection request cannot be divided into several lower speed connection requests and routed separately from the source to the destination. The data traffic on a connection request should always follow the same route. The transceivers in a network node are fixed, hence the wavelength continuity constraint applies. In the following formulation, we assume that each grooming node has unlimited multiplexing and demultiplexing capability. This means that the network node can multiplex/demultiplex as many low-speed traffic streams to a lightpath as needed, as long as the aggregated traffic does not exceed the lightpath capacity. However, in practical systems, it may be a constraint.

The following additional parameters are needed for formulation over and above that defined in the last section.

C_l : Cost of using link l .

Variable $p_{l,w}^i$: =1 if wavelength w on link l carries primary, else 0.

Variable $r_{l,w}^i$: =1 if wavelength w on link l carries backup, else 0.

W_l : Number of wavelengths required on link l .

$M_{l,w}$: Primary capacity reserved on wavelength w on link l .

$R_{l,w}$: Backup capacity reserved on wavelength w on link l .

ILP Formulation I: Shared Backup Multiplexing

1. Objective: To minimize the total wavelength links with cost $C_l = 1$.

$$\min \sum_{l \in E} C_l \times W_l \quad (10.17)$$

2. Constraints on physical route variables:

$$p_{l,w}^i = \sum_{p=1}^{KW} \delta^{i,p} \epsilon_l^{i,p} \psi_w^{i,p} \quad r_{ij,w}^m = \sum_{r=1}^{KW} \nu^{i,r} \epsilon_l^{i,r} \psi_w^{i,r} \quad (10.18)$$

3. Constraints on path indicators: Each request gets one primary (backup)

$$\sum_{p=1}^{KW} \delta^{i,p} = 1 \quad \sum_{r=1}^{KW} \nu^{i,r} = 1 \quad (10.19)$$

4. *Constraints on topology diversity*: Disjoint primary and backup paths.

$$\sum_{p=1}^W \delta^{i,p} = \sum_{r=W+1}^{KW} \nu^{i,r} \quad \sum_{p=W+1}^{KW} \delta^{i,p} = \sum_{r=1}^W \nu^{i,r} \quad (10.20)$$

5. *Constraints on wavelength capacity*: Wavelength capacities should not exceed.

$$M_{l,w} = \sum_i d_i \times p_{l,w}^i \quad M_{l,w} + R_{l,w} \leq C \quad (10.21)$$

6. *Constraints on fiber capacity*: $x_{l,w}$ is the sum of primary and backup paths that use wavelength w on link l . $u_{l,w} = 1$, if $x_{l,w} \geq 1$, and zero otherwise.

$$x_{l,w} = \sum_i (r_{l,w}^i + p_{l,w}^i) \quad u_{l,w} \leq x_{l,w} \quad KN(N-1)u_{l,w} \geq x_{l,w} \quad (10.22)$$

$$u_{l,w} \in \{0, 1\} \quad W_l \geq \sum_w u_{l,w} \quad W_l \leq W \quad (10.23)$$

7. *Constraints on backup multiplexing*: The capacity reserved for backup paths accounts for correlation between the corresponding primary paths to determine the maximum value. In the following $j \geq i$.

$$\begin{aligned} R_{l,w} \geq & d_i \times \nu^{i,p} \epsilon_l^{i,p} \psi_w^{i,p} \\ & + \sum_j d_n \times \nu^{j,p,i,p} \epsilon_l^{j,p} \psi_w^{j,p} \times I_{(i,\bar{p}), (j,\bar{p})} + \sum_j d_j \times \nu^{j,\bar{p},i,p} \epsilon_l^{j,\bar{p}} \psi_w^{j,\bar{p}} \times I_{(i,p), (j,\bar{p})} \\ & + \sum_j d_j \times \nu^{j,p,i,\bar{p}} \epsilon_l^{j,p} \psi_w^{j,p} \times I_{(i,\bar{p}), (j,p)} + \sum_j d_j \times \nu^{j,\bar{p},i,\bar{p}} \epsilon_l^{j,\bar{p}} \psi_w^{j,\bar{p}} \times I_{(i,p), (j,p)} \end{aligned} \quad (10.24)$$

where $\nu^{j,p,i,p}$ is a binary variable which takes the value of one when $\nu^{j,p} = 1$ and $\nu^{i,p} = 1$.

$$\nu^{j,p,i,p} \geq \nu^{j,p} + \nu^{i,p} - 1; \quad \nu^{j,p,i,p} \leq \nu^{j,p}; \quad \nu^{j,p,i,p} \leq \nu^{i,p} \quad (10.25)$$

ILP Formulation II: Dedicated Backup with MLPS

For this we change the objective function to minimize the total wavelength-links as well as total link-primary-sharing. Let C_{share}^l be the weight of s_l . Most constraints remain the same and a few change as described below.

1. Objective:

$$\min \left(\sum_{l \in E} C_l \times W_l + C_{share}^l \times s_l \right). \quad (10.26)$$

2. *Constraints on backup capacity*: Backup capacities are simply aggregated.

$$R_{l,w} = \sum_i d_i \times r_{l,w}^i \quad (10.27)$$

3. Constraints on link-primary-sharing:

$$s_l \geq \sum_i \sum_w p_{l,w}^i - 1 \quad (10.28)$$

$$s_l \leq \sum_i \sum_w p_{l,w}^i \quad (10.29)$$

10.4.4 Performance Results

The performance of grooming depends on the efficiency of *grooming fractional wavelength traffic* and the *traffic pattern* offered, and on the sequence it is offered. When traffic granularity is close to full-wavelength capacity, grooming cannot bring much improvement on wavelength utilization. Therefore, we generate traffic randomly with each request requiring 1/4 capacity of a full wavelength. Two link disjoint paths for each connection are pre-computed using the shortest-path routing algorithm. The experiments were performed on a 10-node 14-bidirectional link topology as depicted in Fig. 10.6 (right).

In one experiment with 23 requests, the formulation I (backup multiplexing) required 28 wavelength-links, while formulation II (dedicated backup) required 33 wavelength-links. In general, formulation II requires more wavelength-links than formulation I. However, in further experimentation, we also noticed that dedicated protection becomes affordable under heavy traffic scenarios as the wavelength utilization significantly improves with traffic grooming. Also formulation II is computationally less expensive and hence more practical.

10.4.5 Partial Protection

Our approaches of survivable grooming network design can be extended to the partial protection in WDM grooming networks. The grooming capability of the network makes partial protection a possible solution when the network resources are not sufficient to provide full protection for every request. It may be a preferred solution as failures are not that often and a graceful degradation of services may be acceptable when compared with no service at all even in the failure-free case. Partial protection is defined as follows.

For a request m , its requested capacity for primary or working path is given as d_m , the minimum capacity for its backup is given as b_m . The difference between partial protection and full protection is that here $0 < b_m < d_m$, while in the full protection, $b_m = d_m$. Just for the sake of completeness, when $b_m = 0$, it is called no protection for request m . The problem of partial protection is to find a primary path for request m , assigning capacity of d_m to it, and find a backup path with capacity c_m such that $b_m \leq c_m \leq d_m$. The higher the value of c_m , the better the protection request m has.

The exact ILP formulations can be modified to solve the partial protection problems in grooming networks as well. However, a direct modification makes the formulations nonlinear, because in partial protection problems, the backup capacity becomes unknown. If we reconsider the motivation of the partial protection in grooming networks, the problem might be solved differently. The main reason partial protection is adopted is that we do not have enough wavelength resource to provide full protection for each request. In other words, we may not want to exploit one extra wavelength just to provide more than the minimum capacity requirement of the backups. In this situation, the partial protection problem can be divided into two subproblems.

1. Resource minimization: given the network resource and minimum backup requirement, try to allocate each request m with primary capacity of d_m and backup capacity of b_m .
2. Protection maximization: given that all the requests are accommodated with the minimum protection requirement being satisfied, the second step is to optimally distribute the residual network capacity to provide better protection to some, if not all, of the requests.

Using the above, we propose a two-phase ILP formulation with dedicated backup reservation for the partial protection design in WDM grooming networks as follows. The details are omitted here and can be found in [14]. The experiments are performed on the same 10-node network topology. For a particular traffic matrix, a full protection requires 33 wavelength links whereas, for $P_{ratio} = 0.6$, a total of only 28 wavelength-links are required. We also noticed that many of the requests were provided with more capacity than their minimum requirements and some connection requests are fully protected. In another experiment, in the same topology with each link consisting of one fiber that carries 3 wavelengths, we experimented to evaluate blocking probability. It is assumed that random requests arrive at each node according to a Poisson process with rate λ . Each request is equally likely to be destined to any of the remaining nodes. The holding time of the requests are exponentially distributed with mean $1/\mu$. The requested capacity is uniformly distributed between a given lower- and an upper-bound, and the full wavelength capacity is chosen to be OC-48. A request is said to be accepted if and only if both a primary path and a backup path for the request can be successfully allocated.

We performed experiments with two types of traffic. In one case, the request capacity is uniformly distributed between OC-1 and OC-36. In the other case, the request capacity is uniformly distributed between OC-24 and OC-36. Fig. 10.7 presents the networking blocking performance as the node load changes. For each node load, we perform simulations in 10 rounds, with each round having 100,000 random requests. An average value is taken as the blocking probability for the given node load value. It is observed from Fig. 10.7 that as the protection ratio goes down, the network blocking performance improves. This is as expected. More importantly, there is a significant gain in blocking performance.

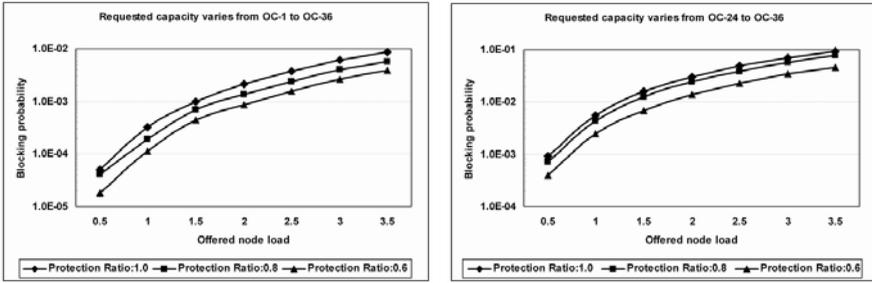


Fig. 10.7. Blocking performance for traffic (1–36) and (24–36)

10.5 Conclusion

A major issue in optical fiber network is the management of fault due to the huge amount of traffic carried by a single fiber. Capacity efficiency and recovery speed are two important aspects in designing protection mechanisms. We presented the resource planning where a single link failure is part of the design and operation process. Shared backup path protection, its variations, and p -Cycle protection are the most widely studied protection schemes in mesh network protection. Shared-path protection are believed to be capacity efficient. Studies have shown that general shared-path protection and pre-cross-connected shared-path protection use less total capacity than p -Cycle protection in low-connectivity networks, while they are comparable in high-connectivity networks.

The protection and restoration design in grooming networks is more complicated than that of the conventional WDM networks. Although shared wavelength protection is still the preferred option, our study shows that dedicated backup reservation becomes affordable and appears to be more desired when the wavelength utilization is improved by the grooming capability of the network. Furthermore, by adding a constraint to minimize the total link-primary-sharing, the number of affected working paths due to single link failure is reduced. A good practical example of designing both traffic grooming and survivability in practice is a Light-trail architecture as discussed in [18].

Dynamic establishment of restorable connections is another important issue. Due to computational complexity of the problem, developing practical and efficient heuristic approaches for survivable design remains an important research direction. When network resource is restrained and insufficient to provide 100% protection to every request, one solution is to provide partial protections. Our results show that partial protection is an effective compromise when the network resources are limited.

Acknowledgements The research is in part supported by Jerry R. Junking endowment at Iowa State University and National Science Foundation

grants CNS0323374, CNS0439748, and CNS0626741. The detailed work is reported in PhD dissertations of Drs. J. Fang [19], W. He [20], and S. Balasubramanian [21].

References

1. B. T. Doshi, S. Dravida, P. Harshavardhana, O. Hauser, and Y. Wang, "Optical Network Design and Restoration", *Bell Labs Technology Journal*, pp. 58–84, January 1999.
2. S. Ramamurthy and B. Mukherjee, "Survivable WDM mesh networks, part I: protection," *IEEE INFOCOM*, vol. 2, pp. 744–751, March 1999.
3. S. Ramamurthy and B. Mukherjee, "Survivable WDM mesh networks, part II – restoration", *IEEE ICC*, vol. 3, pp. 2023–2030, June 1999.
4. M. T. Frederick, P. Datta and A. K. Somani, "Sub-Graph Routing: A generalized fault-tolerant strategy for link failures in WDM Optical Networks," *Computer Networks* 50(2): 181–199, 2006.
5. G. Ellinas, G. Halemariam, and T. Stern, "Protection cycles in mesh WDM networks", *IEEE JSAC*, vol. 18, no. 10, pp. 1924–37, October 2000.
6. D. A. Schupke, C. G. Gruber, and A. Autenrieth, "Optimal configuration of p-cycles in WDM network", *Proceeding of IEEE ICC*, pp. 2761–2765, April 2002.
7. J. Doucette, D. He, W. D. Grover, and O. Yang, "Algorithmic approaches for efficient enumeration of candidate p-cycles and capacitated p-cycle network design," *Proceeding of DRCN 2003*, pp. 212–220, October 2003.
8. N. Jose and A. K. Somani, "Connection Rerouting/Network Reconfiguration," in Proc. of 4th DRCN 2003, Banff, Alberta, Canada, October 2003, pp. 23–30.
9. O. Gerstel and R. Ramaswami, "Optical network survivability: A service perspective", *IEEE Communication Magazine*, vol. 38, no. 3, pp. 104–113, March 2000.
10. M. Sridharan, A.K. Somani, and M.V. Salapaka, "Approaches for capacity and revenue optimization in survivable WDM networks," *Journal of High Speed Networks*, vol. 10, no. 2, pp. 109–125, August 2001.
11. G. Mohan, C. S. R. Murthy, and A. K. Somani, "Efficient algorithms for routing dependable connections in WDM optical networks," *IEEE/ACM Transactions on Networking*, vol. 9, no. 5, pp. 553–566, October 2001.
12. G. Shen and W. D. Grover, "Performance of protected working capacity envelopes based on p-cycles: Fast, simple, and scalable dynamic service provisioning of survivable services," in *Proceeding of APOC*, vol. 5626, November 2004.
13. S. Thiagarajan and A. K. Somani, "Traffic Grooming for Survivable WDM Mesh Networks," in Special Issue on Protection/Restoration Meets the Reliability Challenge, *Optical Networks Magazine*, pp. 88–98 May/June 2002.
14. J. Fang, M. J. Fang, M. Sivakumar, A. K. Somani and K. M. Sivalingam, "On Partial Protection in Groomed Optical WDM Mesh Networks," in Proc. of IEEE Dependable Systems and Networks – Dependable Computing and Communications Symposium, Japan, pp. 228–237, June 2005.

15. W. He, J. Fang and A. K. Somani, "A new p-cycle based Survivable Design for Dynamic Traffic in WDM Networks," in Proc. of IEEE Globecom 2005, November 2005.
16. A. K. Somani, "Survivability and Traffic Grooming in Optical Networks," Cambridge Press, January 2006, 436pp.
17. C. Ou, K. Zhu, H. Zhang, L. H. Shastrabuddhe, and B. Mukherjee, "Traffic grooming for survivable WDM networks – shared protection," *IEEE Journal on Selected Areas in Communications*, pp. 1367–1383, November 2003.
18. J. Fang, W. He and A. K. Somani, "Optimal light trail design in WDM optical networks," *IEEE ICC 2004*, Vol. 3, pp. 1699–1703, June 2004.
19. J. Fang, "Traffic grooming in IP over WDM optical networks," PhD thesis, Iowa State University, 2004.
20. W. He, "Survivable design in WDM mesh networks," PhD thesis, Iowa State University, 2006.
21. S. Balasubramaniam, "Design and Protection Algorithms for Path Level Aggregation of Traffic in WDM Metro Optical Networks," PhD thesis, ISU, 2007.

Traffic Grooming Under Scheduled Service

Bin Wang

11.1 Introduction

Traffic grooming offers the ability to switch low speed traffic streams into high speed bandwidth trunks so that resources are shared among multiple entities that individually need only a fraction of the resources, which optimizes the capacity utilization by allowing the resource cost to be amortized over the number of users [1, 2]. Although much of today's physical layer network infrastructure (e.g., metro networks) have been built with ring topologies using add/drop multiplexers to add/drop wavelengths and/or tributary circuits, WDM optical networks are evolving toward general mesh topologies. Indeed, much recent work has focused on grooming traffic in mesh networks. Existing work on traffic grooming has considered several types of traffic model, such as static traffic model, dynamic random traffic model, incremental traffic model, traffic matrix set model, and scheduled traffic model. In the static traffic model, all traffic demands are known in advance and do not change over time. For instance, a client company may request virtual private network capacity for connectivity among different company sites from a service provider. The objective is typically to minimize the network resources needed or cost, e.g., the amount of line terminating equipment (LTE), or to maximize the network throughput given a resource constraint, and so on. This model does not allow dynamic connection setup and tear-down. In the dynamic random traffic model, a demand is assumed to arrive at a random time and last for a random amount of time. Usually this model assumes a certain stochastic demand arrival process (e.g., Poisson process) and a demand holding time probability distribution (e.g., exponential distribution), as well as a certain spatial traffic distribution (e.g., uniform traffic). The design objective is typically to minimize the percentage of blocked traffic. Other traffic models have been considered for network planning and configuration. The work in [3] on multi-period network planning was based on an incremental traffic model and conducted network planning across several years to incrementally produce a network capable of carrying all traffic predicted up to the end of the planning

horizon. The work reported in [4] considered time-varying offered traffic in the form of a set of traffic matrices at different instants for off-line configuration so as to accommodate the time-varying traffic. The work in [5] also used a set of traffic matrices to design and dimension a WDM mesh network to groom dynamic traffic.

While these different traffic models are valid and useful in many circumstances, they are not able to capture the traffic characteristics of applications that require capacity during a specific time interval or for a certain time period. For instance, a client company may request some bandwidth from a service provider to satisfy its communication requirements at a specific time, e.g., between headquarters and production centers during office hours or between data centers during the night when backup of databases is performed. Another example is that an IP service provider may use leased static lightpaths as IP links for providing guaranteed network capacity at all times. However, during peak traffic hours (e.g., working hours of 8 am to 5 pm), the IP network needs additional links. The additional links can be realized by periodic lightpaths that are scheduled from 8 am to 5 pm daily. Other examples include many large-scale science applications (e.g., applications in high energy physics, climate simulation, astrophysics, remote control of scientific instruments, etc.) that must deliver, at scheduled time durations, hundreds of gigabits per second throughput in near future and several terabits per second within the next decade, ranging from cooperative remote visualization of massive archival data through the distribution of large amounts of simulation data, to the interactive evolution of computations through computational steering [6]. These applications require the provisioning of scheduled dedicated channels or sub-wavelength bandwidth pipes at a specific time with a certain duration. Consequently, traffic models that capture the time-limited nature of scheduled services have been proposed and studied recently. New on-demand, time-limited large-bandwidth services are beginning to be offered by operators [7]. Traffic grooming under scheduled service is essential because many applications only need low-rate connections for communication, for instance, connections between corporate sites, scheduled backup between data centers, and occupying a full wavelength for transferring a few megabits or gigabits per second of data results in very poor utilization of network resources.

Along with the scheduled traffic models, a host of issues and opportunities arise in the study of traffic grooming for network planning purpose and for service provisioning as well. In this chapter, we will survey different traffic models that capture various aspects of scheduled service, summarize some recent work conducted on traffic grooming under these models, and discuss some related issues for future research.

The rest of the chapter is organized as follows. Section 11.2 surveys traffic models for scheduled service. Recent work on traffic grooming under scheduled service is introduced in Section 11.3. Section 11.4 then discusses some related issues and future research directions of grooming under scheduled service. Section 11.5 summarizes the chapter.

11.2 Scheduled Service Traffic Model

Several forms of traffic models [8, 9, 10, 11, 12, 13] have been considered over the years to capture various aspects of scheduled service. A demand is typically characterized by a source, a destination, and a bandwidth requirement. Depending on the starting time and the holding time of the demand, four types of scheduled service are summarized as follows.

11.2.1 Demands with Known Starting Time and Known Holding Time

This type of demands has a known starting time and a known holding time. The requested bandwidth needs to be provisioned at the starting time and the connection will be torn down after the holding time ends. A specific example is the scheduled lightpath demand (SLD) model [8, 14, 15]. The objective of service provisioning is typically to find a feasible route in the network that minimizes the network resources used or maximizes the likelihood of accommodating future demands, subject to network physical constraints, such as link capacity constraint, wavelength continuity constraint, and so on. This traffic model captures many features of aforementioned applications and has been studied by many researchers [8, 15, 16, 17, 18, 19, 20, 21, 22].

11.2.2 Demands with Known Starting Time and Unknown Holding Time

This type of demands has a known starting time and its holding time is not specified. The objective of service provisioning is typically to find a feasible route in the network in advance to increase the likelihood of successful reservation. Additionally, the network resources used to accommodate the demand should be minimized. Sometimes, the holding time can be capped by a maximum time limit. This traffic model is useful for characterizing relatively long lasting future demands that expect to get as long network services as possible. This model is suitable for providing communication service for applications that are of low priority and try to take advantage of idle network resources as much as possible.

11.2.3 Demands with Unknown Starting Time and Known Holding Time

The exact starting time of this type of demands may be unknown. The holding time is, however, pre-specified. In its simple form, this type of demands can implicitly assume a starting time of current time when the demand arrives and is appropriate for characterizing dynamic demands with known holding times. In general, the demands need to be provisioned as early as possible.

This traffic model applies to applications that expect to get network services as early as possible, but can tolerate a service delay. This model is suitable for providing communication service for non-critical data backup.

11.2.4 Demands with Known Holding Time and Scheduling Time Flexibility

Many applications have a certain degree of flexibility about when exactly the scheduled demand should occur. Service providers can therefore exploit the flexibility to schedule the connections to achieve a better utilization of resources and reduce cost.

One example of this type of traffic model is the sliding scheduled traffic model proposed in [13]. A sliding scheduled demand is represented by a tuple $d = (s, t, b, \ell, r, h)$ that satisfies $r - \ell \geq h > 0$, where s and t are the source and destination, b is the number of requested capacity units, and ℓ and r are the starting time and ending time of a time window during which the demand with a holding time of h time units resides. In this model, the demand holding time h is an interval within a time window $[\ell, r]$. Rather than fixing the starting time and ending time of the demand, a scheduling time flexibility is introduced with this model. As a result, the demand is allowed to slide within a larger time window $[\ell, r]$. This model allows an application to specify a larger time window during which the demand for communication capacity is satisfied. This model gives a service provider more flexibility in provisioning the requested demand and a better opportunity to optimize the network resources since a demand is considered accommodated as long as it is provisioned within the larger time window. Given a demand d , the actual starting time of the demand is variable relative to the left boundary ℓ of its associated time window. If the demand starts at a time units after ℓ , the demand is active during $[\ell + a, \ell + a + h]$.

This model can be considered as a generalization of the three types of traffic model defined earlier. For example, a demand with known starting time and known holding time is a special case when the demand starting time is the starting time of the time window while the holding time is the same as the window size. A demand with known starting time and unknown holding time is also a special case in that the time window is unbounded and the holding time is the same as the time window size. The third model is yet another special case where the time window size is infinitely large.

The work of [23] considers a similar traffic model where the scheduling of periodic connections with time flexibility on a single WDM link is studied.

11.3 Traffic Grooming under Scheduled Service

The work on traffic grooming under scheduled service is more recent and relatively limited compared to work on traffic grooming in general. In this section, we introduce some recent endeavors in this area.

11.3.1 Dynamic Traffic Grooming of Subwavelength Connections with Known Duration

The work by Tornatore et al. [21] considers dynamic traffic grooming of sub-wavelength connections with known duration where the connection requests arrive one at a time with different starting time and holding time. The objective of grooming is to minimize the network resources used for accommodating each request or implicitly attempt to minimize the overall connection blocking probability.

The authors adopt an auxiliary graph (AG) $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ to expand the representation of the original network where \mathcal{V} and \mathcal{E} are the set of nodes and links, respectively. Assuming that each link has W wavelengths, the auxiliary graph \mathcal{G} is a layered graph with $W + 2$ layers where layers 1 to W denote the W wavelength layers, layer $(W + 1)$ is the lightpath layer where established lightpaths are represented, and layer $(W + 2)$ is the access layer where a traffic demand starts and terminates. When a lightpath is established, the wavelength-links used are deleted in the appropriate wavelength layers of the graph. Similarly, corresponding wavelength-links are restored in the wavelength layers when a lightpath terminates. Network resources such as transceivers, converters, wavelength-links, and lightpaths, can be modeled using proper links in the layered auxiliary graph. The cost of links is properly assigned and updated to reflect the cost of each network element, such as transceiver, converters, wavelength-link, etc. An edge is also characterized by a proper capacity. For example, a wavelength-link edge has a capacity of a wavelength whereas for a lightpath edge, its capacity is the residual capacity after accommodating traffic demands. By manipulating the cost and capacity of links, the auxiliary graph approach allows different grooming policies to be applied while taking into account various network constraints such as transceiver availability, wavelength conversion capabilities, and grooming capabilities.

Given that the connection duration information is known upon the arrival of a demand, the authors design a holding-time-aware provisioning algorithm (HTA). The connection duration information allows the residual holding time of all the existing lightpaths in the network to be evaluated. Because a lightpath usually supports multiple time-limited connections, the residual holding time of a lightpath is determined by the connection(s) with the maximum remaining holding time after which the lightpath can be terminated. With respect to the new demand, an existing lightpath, if chosen, can serve the demand if its holding time is smaller than the residual holding time of the lightpath and the remaining capacity is sufficient to accommodate the demand. If the residual holding time of the lightpath is smaller than the demand holding time, the lifetime of the lightpath may need to be prolonged in order to accommodate the demand. An efficient routing strategy is then designed based on a holding-time-aware edge cost assignment in the auxiliary graph. The idea is to assign a lower cost to an existing lightpath whose lifetime

does not need to be prolonged, and therefore encourage the routing algorithm to reuse the lightpath for grooming. Start from a network represented as an AG with uniform cost p_e for any type of link. Given a new request (s, t, b, h) where s and t are the source and destination, b is the capacity requirement, and h is the holding time, the cost of an edge e is updated using the following cost function C_e where ϵ is a small constant:

$$C_e = \begin{cases} p_e \epsilon & \text{if } e \text{ is a lightpath with residual holding time } H_r \geq h, \\ p_e \epsilon + p_e \Delta t & \text{if } e \text{ is a lightpath with residual holding time } H_r < h \\ & (\Delta t = h - H_r), \\ p_e h & \text{for all other wavelength-links in AG.} \end{cases} \quad (11.1)$$

A shortest path algorithm is run on the auxiliary layered graph with respect to the assigned link cost to groom the new demand. Once the demand is groomed, the links, link costs, and capacity will be updated accordingly.

Simulation results [21] reported by the authors indicate a significant advantage in terms of reducing blocking probability by employing the holding-time-aware grooming approach as opposed to other efficient approaches that are holding-time unaware.

11.3.2 Traffic Grooming under a Sliding Scheduled Service Model

In [12], the authors study traffic grooming under a sliding scheduled service model. No wavelength conversion is assumed in the network under consideration. And no traffic bifurcation is allowed. Given a set of sliding-scheduled traffic demands, the primary objective of traffic grooming is to minimize the total wavelength-links used while trying to meet demands' timing specifications (i.e., starting time and holding time). Because of the scheduling time flexibility offered by the sliding scheduled service model, the difficulties of the traffic grooming problem lies in the spatial and temporal constraints/flexibilities imposed/offered on the set of demands. In the spatial domain, demands are groomed in a mesh network topology and may share the same wavelength on the same link when the capacity of the wavelength allows. Grooming of demands is also subjected to the wavelength continuity constraint. In the time domain, a demand may slide within its time window and demands may overlap in time, which constrains traffic grooming when combined with the spatial constraints. However, the time-limited nature of demands offers additional opportunities for resource reuse in both time and space during traffic grooming. Obviously, reducing overlapping between demands in time helps temporal resource reuse.

A two-step approach to the traffic grooming problem is proposed and studied. First, to maximize the network resource reuse by demands in the time domain, a demand time conflict reduction algorithm is designed. Given a set of demands \mathcal{D} , the time domain conflict reduction algorithm is applied

to \mathcal{D} to find a proper placement of demand intervals in their associated time windows such that the number of demand pairs that overlap in time is minimized. This problem is solved by first constructing an interval graph \mathcal{H} based on the demands' time windows [13]. Then strong and weak edges in the graph \mathcal{H} are identified [13]. Note that strong edges reflect time conflicts of demand pairs that cannot be resolved no matter how demands are placed within their time windows whereas weak edges represent time overlapping between two demands that can be avoided if they are properly scheduled within their time windows. The proposed time conflict reduction algorithm works to remove as many weak edges as possible from the graph \mathcal{H} to obtain an edge-reduced graph \mathcal{H}' . Therefore, strong edges of \mathcal{H} are always in the resulting graph \mathcal{H}' . However, whether a weak edge of \mathcal{H} is in \mathcal{H}' depends on how demands are placed in their associated time windows.

Once demands are properly placed in their time windows to obtain a new demand set \mathcal{D}' (i.e., with demands' start time fixed), the second step is to use a time window based algorithm for traffic grooming. The design of this grooming algorithm is based on the following observations. Demands that overlap in time must be groomed with sufficient capacity in the spatial domain if the total capacity of these demands exceeds the wavelength capacity. Network resources used by one demand can be reused by another demand as long as they do not overlap in time. This motivates the authors to divide the set of demands into subsets based on the demands' starting time and ending time such that demands in different subsets are disjoint in time except for a few straddling demands. Note that it is not always possible to assign a demand to only one subset because its holding time may result in its being included in multiple subsets. Such a demand is termed as a straddling demand. Specifically, the time conflicts of demands in \mathcal{D}' are modeled using a different interval graph. Based on the interval graph representation, an efficient algorithm then divides the demands into subsets. A virtual network topology is constructed to groom each subset of demands. The resources used by demands in one subset can potentially be reused by other subsets. Grooming of a straddling demand, however, requires the same resources to be used by it across subsets.

Simulation results and comparison with a customized tabu search scheme have shown that the proposed two-step traffic grooming algorithm for sliding scheduled demands is effective in meeting demands' timing specifications as well as reducing the total network resources used [12].

11.3.3 Traffic Grooming under a Scheduled Service Model Using Light-Trails

Our recent work considers the problem of accommodating a set of subwavelength scheduled traffic demands in a shared wavelength optical network [24] (e.g., a light-trail network) with the objective of minimizing the total resources used. Specifically, consider a network $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ where \mathcal{V} is the set of nodes and \mathcal{E} is the set of links in the network. Each link has W wavelengths and a

wavelength has a capacity of C . We assume no wavelength conversion capability in the network. Let $\mathcal{D} = \{d_1, d_2, \dots, d_m\}$ be the set of scheduled traffic demands. A demand d_i is represented as $(s_i, t_i, b_i, \alpha_i, \beta_i)$ where s_i is the source node, t_i is the destination node, b_i is the requested capacity, and α_i and β_i are the starting time and the ending time of the demand. The problem is to construct a set of light-trails $\mathcal{LT} = \{lt_1, lt_2, \dots, lt_h\}$ such that all the demands in \mathcal{D} are accommodated with the objective of minimizing the total resources used, such as the number of wavelength-links used, the number of wavelengths used.

When constructing light-trails to accommodate the scheduled demands, we consider two cases: *static light-trail case* and *dynamic light-trail case*. In the static light-trail case, a light-trail will not be terminated once it is set up. In the dynamic light-trail case, a light-trail can be torn down when it is not used by any scheduled demand, so that the resources for that light-trail can be re-allocated for other light-trails. In the dynamic light-trail case, the light-trails are time limited. Integer linear programming formulations for both cases are derived. Due to the complexity of solving ILPs for large problems, heuristic iterative grooming algorithms are proposed.

Heuristic Algorithm: Static Light-Trail Case

The proposed algorithm is shown in Algorithm 1 which combines a greedy packing algorithm with evolutionary iterations. In *Heuristic_Light-Trail*, k shortest paths are generated for the source–destination pair of each demand and are included in Γ as candidate light-trails. Then light-trails are selected using *ResAlloc* (not shown) from Γ and instantiated so that every demand in the demand set is accommodated by one of the light-trails and the total cost is minimized. *ResAlloc* returns (Γ_{ins}, c) , where Γ_{ins} is the set of instantiated light-trails and c is the total cost.

The solution is improved iteratively. In each iteration, the algorithm removes a candidate light-trail lt ($lt \in \Gamma_{ins}$) from the candidate set Γ and assigns the rest of the light-trails in Γ to a new set Γ' : ($\Gamma' \leftarrow \Gamma \setminus lt$). If by using Γ' as the candidate light-trail set, algorithm *ResAlloc* produces a better solution (i.e., with a lower cost), the current best solution is replaced. The criteria of improvement is whether the cost of the current solution is reduced enough from the previous best solution by a tunable threshold θ (e.g., 0.5%). In the case where each of the light-trails in Γ_{ins} has been tried but no improvement is obtained, the algorithm breaks out from the iteration and returns the current best solution obtained. Otherwise, the iteration continues.

The idea behind *ResAlloc* is explained briefly. A candidate light-trail can carry a demand only if both the source and destination nodes of the demand are on the light-trail and the source node is on the upstream of the destination node. A demand that can be carried by a light-trail is *supportable* by that light-trail. The algorithm sorts all the demands in the demand set that are *supportable* by a candidate light-trail lt in a certain order. Different ordering schemes are possible. Studies have shown that the longest span first (LSF) is

Algorithm 1 *Heuristic_Light-Trail*(\mathcal{G}, \mathcal{D})

Require: \mathcal{G} is the network topology and \mathcal{D} is the set of scheduled demands.

```

1:  $(\Gamma_{ins}, c_0) \leftarrow (\emptyset, |\mathcal{E}| \cdot W)$ ,  $flag \leftarrow \text{TRUE}$ ;
2: Find  $k$ -shortest paths for each demand  $d_i \in \mathcal{D}$  and put them in  $\Gamma$  as the set of
   candidate light-trails;
3:  $(\Gamma_{ins}, c_0) \leftarrow ResAlloc(\mathcal{G}, \Gamma, \mathcal{D})$ ;
4: while  $flag == \text{TRUE}$  do
5:    $flag \leftarrow \text{FALSE}$ ;
6:   for all  $lt$  such that  $lt \in \Gamma_{ins}$  do
7:      $\Gamma' \leftarrow \Gamma \setminus lt$ ;
8:      $(\Gamma_{temp}, c) \leftarrow ResAlloc(\mathcal{G}, \Gamma', \mathcal{D})$ ;
9:     if  $(c_0 - c)/c_0 \geq \theta$  then
10:       $c_0 \leftarrow c$ ;
11:       $\Gamma_{ins} \leftarrow \Gamma_{temp}$ ;
12:       $flag \leftarrow \text{TRUE}$ ;
13:     end if
14:   end for
15: end while

```

most effective [25]. The span of a demand is defined as the number of hops between the source node and the destination node of the demand. The LSF scheme attempts to better utilize the space (span) of the light-trail. A demand with a span of n hops on a light-trail with a length of m ($m \geq n$) would leave other $(m - n)$ links on the light-trail idle when the demand is transmitting data. Therefore, we prefer to select the demand with a longer span when packing demands onto the light-trail so that there are fewer links being idle. If two demands have equal spans, the one with a larger requested capacity is selected first. These demands are packed on the light-trail in the same order until the wavelength capacity limit is reached or all the *supportable* demands of that light-trail have been packed.

Once all the candidate light-trails are packed with supportable demands, the utilization ratio of a light-trail lt , ρ_{lt} , is calculated as:

$$\rho_{lt} = \sum_{\forall d_i \text{ packed onto } lt} b_i \cdot (\beta_i - \alpha_i) / (C \cdot \Psi), \quad (11.2)$$

where C is the wavelength capacity and $\Psi = [|\text{Min}(\alpha_i), \text{Max}(\beta_i)|]$ and **Min**, **Max** are taken over demands packed onto light-trail lt . The algorithm then selects the candidate light-trail lt_b with the highest utilization ratio, and moves it to the set Γ_{ins} to be instantiated. All the demands that are packed onto this light-trail are removed from the demand set \mathcal{D} . Next, each remaining candidate light-trail is packed again with the demands that are left in \mathcal{D} . Based on the utilization ratio another best candidate light-trail is picked. The iteration repeats until there is no demand left in \mathcal{D} . The algorithm then assigns wavelengths to the light-trails in Γ_{ins} and calculates the total cost.

Heuristic Algorithm: Dynamic Light-Trail Case

The basic idea behind the algorithm for the dynamic case is similar to the static light-trail case. The heuristic algorithm shares the same main function of *Heuristic_Light-Trail* (Algorithm 1) with *ResAlloc* replaced by a new *ResAlloc* (not shown). A light-trail can be terminated when it is not used by any scheduled demand, so that the resources for that light-trail can be re-allocated for other light-trails. Each light-trail can therefore have several time windows during which it is active.

The question is how to determine whether a light-trail should be made active at a certain time instant. A simple answer depends on whether there is any demand being carried on the light-trail at that time instant. This approach may produce light-trails with a low capacity utilization ratio since a demand with a very small requested capacity will keep a light-trail active. Instead, we can make the light-trail inactive when it has a low capacity utilization ratio and release the resources to other light-trails to improve the resource utilization. In Fig. 11.1(a), it is shown that there are four demands packed onto the light-trail. Each demand is represented by a rectangle in the figure. The height of the rectangle indicates the requested capacity and the width indicates the time duration of the scheduled demand.

The capacity utilization of a candidate light-trail packed with demands varies over time, which can be represented by a capacity utilization curve like the one in Fig. 11.1(b). We introduce a capacity threshold σ and deem a light-trail active at a certain time instant if the capacity utilization ratio is no less than σ , or inactive otherwise (Fig. 11.1(c)). Thus a light-trail would be either active or inactive at any time instant, which can be shown by an activation

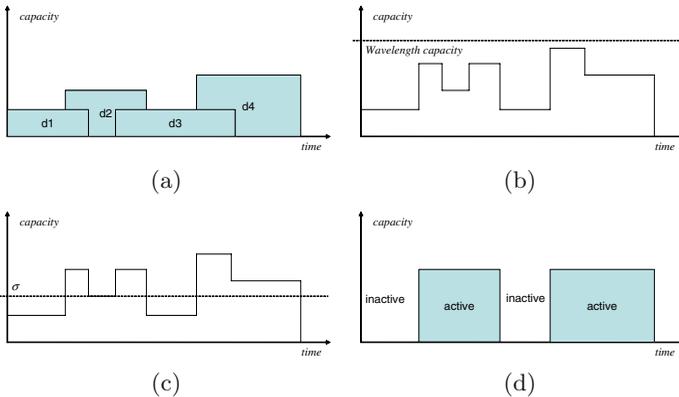


Fig. 11.1. An example of packing a time-limited light-trail with demands: (a) the demands initially packed in a light-trail; (b) the initial capacity utilization curve; (c) the capacity utilization curve with a threshold; (d) the activation curve of the light-trail

curve like the one in Fig. 11.1(d). If a demand packed in the light-trail is active during an inactive period of the light-trail, the demand will be removed from the light-trail unless the demand has the same source and destination as the light-trail, and be packed onto other potential light-trails. The demand will be accommodated by at least one light-trail because k shortest paths are generated for the source–destination pair of each demand and are included in \mathcal{L} as candidate light-trails.

The solutions to the ILP problem formulations and simulation of the heuristic algorithms show that, compared to using static light-trails, the use of dynamic light-trails for traffic grooming significantly reduces the required resources by as much as 30% for a demand set of size 100. The resource reduction becomes more significant as the size of demand sets increases and the time correlation among demands becomes weaker. The heuristic algorithms for both cases are also shown to be very time efficient. Details of simulation study can be found in [25].

11.4 Issues and Future Research Directions

The traffic grooming problem, in general, can be logically decomposed into four subproblems: (1) topology design that determines the virtual topology to be embedded in the physical topology; (2) routing that determines the route of each of the high-speed bandwidth trunks (e.g., lightpaths, light-trails, etc.) over the physical topology; (3) connection routing that routes connections over the virtual topology; and (4) wavelength assignment that allocates wavelengths to the bandwidth trunks subject to assignment and continuity constraints. In this section, we discuss related issues for grooming under scheduled service as well as future research directions.

11.4.1 Virtual Topology Design

Traditionally, designing a virtual topology on a physical network consists of deciding the lightpaths to be set up in terms of their source and destination nodes and wavelength assignment. A related problem is reconfiguring a network from one virtual topology to another. A virtual topology is designed on the basis of traffic pattern and a physical topology. Being able to reconfigure virtual topology provides adaptability, self-healing capability, and upgradability. Two types of approaches have been considered for virtual topology reconfiguration. In the cost approach, the physical network topology, the current virtual topology as well as the new virtual topology that the network must be reconfigured to are known. The goal is to minimize the cost of the reconfiguration, e.g., the cost of optical switching reprogramming, establishment of new lightpaths, and elimination of old ones. The optimization approach assumes that only the current virtual topology is given and changed

traffic pattern and/or physical topology makes reconfiguration necessary. Reconfiguration involves solving a new virtual topology design problem.

Both virtual topology design and reconfiguration have not considered the time-limited nature of scheduled service. The timing information of scheduled service demands offers another dimension for optimization as demonstrated in grooming scheduled service using light-trails (Section 11.3.3) where light-trails can be time-limited. However, significant difficulties exist in exploiting the time-limited nature of scheduled service for the betterment of traffic grooming, especially in case of dynamic traffic arrivals. The optimal virtual topology changes over time. Rather than using the best topology in every time period and incurring a possibly significant reconfiguration cost, it is necessary to consider the optimum sequence of virtual topologies that will minimize the sum of the operating and reconfiguration costs over the entire horizon. Other considerations include virtual topology design for traffic grooming [26] that (1) minimizes network resource (e.g., wavelengths, ports) usage subject to given traffic blocking requirements; and (2) maximizes performance or revenue given physical topology and resource constraints.

11.4.2 Routing

Routing is important for virtual topology design, routing of connections, and providing survivability. Not much has been done to tailor routing for scheduled service. The difficulty has to do with taking into account the connection holding time of scheduled service. The work of [27] considers the routing of holding-time aware demands under a sliding scheduled traffic model to satisfy the bandwidth, starting and holding time requirements of dynamic traffic demands. Given the network topology, current link state information, and a maximal hop count H , this Bellman-Ford flavored routing algorithm finds, for each hop count value $h, 1 \leq h \leq H$, and destination node t , all h -hop feasible path between s and t under the sliding scheduled traffic model. An h -hop feasible path between s and t is a path that satisfies both the demand's bandwidth and timing requirements. This algorithm is useful for routing scheduled connection requests over a virtual topology during traffic grooming.

11.4.3 Survivability

Survivability is a critical aspect of transport networks because of the inherent vulnerability of transmission systems. Existing work deals with survivability service provisioning in the context of scheduled lightpath service.

Kuri et al. [8, 14] address the problem of routing working and protection paths for scheduled lightpath demands in an optical transport network. The problem is formulated as a combinatorial optimization problem where the objective is to minimize the number of wavelength-links required to instantiate the lightpaths. The a priori knowledge of the SLDs' time-disjointness can be

exploited to minimize the amount of wavelength-links required to instantiate the lightpaths needed. Indeed, the same channel can be allocated to multiple lightpaths, provided that they are time-disjoint. A resource-efficient technique called backup multiplexing can be used for minimizing spare channels. In this technique, the same spare channel may serve to protect multiple lightpaths provided that two conditions do not hold simultaneously: the involved lightpaths are overlapping in time and their working paths share at least one common span. A simulated annealing (SA) based algorithm is designed to find approximate solutions to this optimization problem since finding exact solutions is computationally intractable.

The work of Tornatore et al. [20, 28, 29] exploits the connection-holding-time information to dynamically provision shared-path-protected connections. The proposed algorithm, PHOTO, takes advantage of the knowledge of the holding time of connection requests to minimize resource overbuild due to backup capacity and hence achieve resource-usage efficiency. Based on similar ideas, the authors further propose a holding-time-aware, dynamic, connection provisioning algorithm, PHOTO-GSP, to improve sharing of backup resources in segment based protection [19]. The work of Saradhi [30] considers the provisioning of fault-tolerant scheduled lightpath demands based on a two-step optimization that uses a set of pre-computed routes for working and protection paths.

In [16], we study survivable service provisioning with shared path protection under the scheduled traffic model. We consider the static version of the problem where a set of demands is given, and the setup time and tear-down time of a demand are known in advance. We study time efficient approaches to *approximating* the optimal solution to the problem. The proposed approach is based on an iterative survivable routing (ISR) scheme that utilizes a capacity provision matrix and processes demands sequentially using different demand scheduling policies. The objective is to minimize the total network resources (e.g., number of wavelength-links) used by working paths and protection paths of a given set of demands while 100% restorability is guaranteed against any single failure. The additional information on connection holding time is exploited to optimize the network resources jointly in space (i.e., backup resource sharing) and in time (i.e., taking advantage of time-disjointness amongst demands). Since a demand is considered accommodated as long as it is provisioned during its holding time, time disjoint demands (working path and protection path alike) can therefore share network resources. The proposed algorithm is evaluated against solutions obtained by integer linear programming. Simulation results indicate that the proposed ISR algorithm is very time efficient while achieving excellent performance in terms of total network resources used.

Survivable traffic grooming under scheduled service is a research area yet to be explored. A possible solution could be to route scheduled traffic demands with survivability requirements over survivable lightpaths established using the schemes described above.

11.4.4 Waveband Switching

Wavelength switching is used in current light-path networks to set up connections between node pairs. With the increase in the number of wavelengths per fiber, waveband switching (WBS) [31], wherein wavelengths are grouped into bands and switched as a single entity, has been proposed to reduce the cost and control complexity of switching nodes by minimizing the port count. WBS introduces at least one additional constraint: only the traffic carried by a fixed set of wavelengths may be grouped into a band. On the other hand, a light-trail allows the intermediate nodes along a light path to access the wavelength channel, aiming at the reduction of the total number of wavelengths. Both techniques apply traffic grooming on different levels of a WDM network. Innovative solutions that solve the waveband switching and traffic grooming problem using lightpaths or light-trails [32] in an integrated fashion can reduce the port count of multigranular optical crossconnects and the size of digital crossconnects. The solutions can further benefit from taking advantage of timing information of scheduled service. How to efficiently integrate waveband switching with traffic grooming and dynamically provision scheduled high-speed channels or end-to-end connections are open. Although a new switching unit called a waveband-label switched path is defined in GMPLS to expand the underlying provisioning capabilities of traffic grooming and wavebanding in optical networks, focused efforts are needed to investigate the underlying provisioning capabilities with traffic grooming and wavebanding as well as control plane/signaling issues related to GMPLS/ASON on advance reservation for supporting scheduled services.

11.4.5 Other Research Issues

A host of other research issues exist that are related to traffic grooming under scheduled service, such as supporting scheduled multicast service, dynamic provisioning of survivable scheduled demands, and so on. While we have considered minimizing network resource such as wavelength-links or blocking probability as our main performance metrics, other constraints (e.g., port availability, grooming capability, wavelength conversion) and objective functions (e.g., minimize add/drops, OXCs, fibers, transponders, time slots, electronic conversion) can be incorporated in the problems discussed.

11.5 Conclusions

Provisioning scheduled traffic demands has been recognized as an important class of service to be supported in future networks. Along with the scheduled traffic models proposed, a host of issues and opportunities arise in the research of traffic grooming for network planning purpose and for service provisioning as well. In this chapter, we have surveyed different traffic models that capture

various aspects of scheduled service, summarized some recent work on traffic grooming under these models, and discussed some related issues for future research.

References

1. Somani A (2006) *Survivability and Traffic Grooming in WDM Optical Networks*. Cambridge University Press
2. Balasubramanian S, Somani A (2006) On Traffic Grooming Choices for IP over WDM Networks (Invited). In: *Proceedings of BROADNETS*
3. Geary N, Antonopoulos A, Drakopoulos E, O'Reilly J (2001) Analysis of Optimization Issues In Multi-Period DWDM Network Planning. In: *Proceedings of IEEE INFOCOM*
4. Ricciato F, Salsano S, Belmonte A, Listanti M (2002) Off-line Configuration of MPLS over WDM Networks under Time-Varying Offered Traffic. In: *Proceedings of IEEE INFOCOM*
5. Srinivas N, Murthy C (2004) *Photonic Network Communications*, 7:179–191
6. Rao N, Wing W (2003) *Network Provisioning and Protocols for DOE Large-Science Applications*, Report of DOE Workshop on Ultra High-Speed Transport Protocols and Dynamic Provisioning for Large-Scale Science Applications
7. Verizon (2007) <http://www22.verizon.com/wholesale/bandwidth/>
8. Kuri J (2003) *Optimization Problems in WDM Optical Transport Networks with Scheduled Lightpath Demands*. Ph.D. Thesis, ENST Paris
9. Zheng J, Mouftah H (2001) Supporting Advance Reservation in Wavelength-Routed WDM Networks. In: *Proceedings of 10th International Conference on Computer Communications and Networks*
10. Zheng J, Mouftah H (2002) Routing and Wavelength Assignment for Advance Reservation in Wavelength-Routed WDM Optical Networks. In: *Proceedings of IEEE International Conference on Communications*
11. Zheng J, Mouftah H (2003) A Framework for Supporting Advance Reservation Service in GMPLS-Based WDM Networks. In: *Proceedings of IEEE Pacific Rim Conference on Communications, Computers and Signal Processing*
12. Wang B, Li T, Luo X, Fan Y (2004) Traffic Grooming under a Sliding Scheduled Traffic Model in WDM Optical Networks. In: *Proceedings of IEEE Workshop on Traffic Grooming in WDM Networks*
13. Wang B, Li T, Luo X, Fan Y, Xin C (2005) On Service Provisioning Under a Scheduled Traffic Model in Reconfigurable WDM Optical Networks. In: *Proceedings of BROADNETS*
14. Kuri J, Puech N, Gagnaire M (2003) Diverse Routing of Scheduled Lightpath Demands in an Optical Transport Network. In: *Proceedings of Fourth International Workshop on the Design of Reliable Communication Networks (DRCN)*
15. Kuri J, Puech N, Gagnaire M, Dotaro E, Douville R (2003) *IEEE Journal on Selected Areas in Communications*, 21:1231–1240
16. Wang B, Li T (2006) Approximating Optimal Survivable Scheduled Service Provisioning in WDM Optical Networks with Iterative Survivable Routing. In: *Proceedings of BROADNETS*

17. Li T, Wang B, Xin C, Zhang X (2005) On Survivable Service Provisioning in WDM Optical Networks under a Scheduled Traffic Model. In: Proceedings of IEEE Globecom
18. Li T and Wang B (2005) On Optimal Survivability Design in WDM Optical Networks Under a Scheduled Traffic Model. In: Proceedings of the 5th IEEE International Workshop on Design of Reliable Communication Networks (DRCN)
19. Tornatore M, Carcagni M, Ou C, Mukherjee B, Pattavina A (2006) Efficient Shared-Segment Protection Exploiting the Knowledge of Connection Holding Time. In: Proceedings IEEE Globecom
20. Tornatore M, Zhang J, Pattavina A, Mukherjee B (2005) Efficient Shared-Path Protection Exploiting the Knowledge of Connection Holding Time. In: Proceedings ONDM
21. Tornatore M, Baruffaldi A, Zhu H, Mukherjee B, Pattavina A (2007) Dynamic Traffic Grooming of Subwavelength Connections with Known Duration. In: Proceedings of OFC
22. Saradhi C, Gurusamy M (2005) Circular Arc Graph based Algorithms for Routing Scheduled Lightpath Demands in WDM Optical Networks. In: Proceedings of BROADNETS
23. Su W, Sasaki G (2003) Scheduling of Periodic Transfers with Flexibility. In: 41st Annual Allerton Conference on Communication, Control, and Computing
24. Balasubramanian S, Somani A (2006) Traffic Grooming in Statistically Shared Optical Networks. In: Proceedings of 31st IEEE Conference on Local Computer Networks
25. Luo X, Wang B (2007) Service Provisioning Under a Scheduled Traffic Model Using Light-trails in WDM Optical Networks. In: Proceedings of BROADNETS
26. Xin C, Wang B, Cao X, Li J (2006) *Journal of Lightwave Technology*, 24: 2267–2275
27. Wang B, Deshmukh (2005) *IEEE Communications Letters* 9:936–938
28. Tornatore M, Pattavina A, Mukherjee B (2005) Exploiting Connection Holding Time to Improve Resource Efficiency for Dynamic Provisioning in Shared-Path Protection. In: Proceedings OFC
29. Tornatore M, Ou C, Zhang J, Pattavina A, Mukherjee B (2005) *Journal of Lightwave Technology* 23:3138–3146
30. Saradhi C, Wei L, Gurusamy M (2004) Provisioning Fault-Tolerant Scheduled Lightpath Demands in WDM Mesh Networks. In: Proceedings of BROADNETS
31. Cao X (2004) Waveband Switching in Wavelength Division Multiplexing Networks. Ph.D. Dissertation, Department of Computer Science and Engineering, SUNY Buffalo
32. Ye Y, Woesner H, Chlamtac I (2006) *Journal of Optical Networking* 5:701–714

Dynamic Grooming Algorithms

Byrav Ramamurthy

12.1 Introduction

The emergence of Wavelength Division Multiplexing (WDM) technology provides the capability for increasing the bandwidth of Synchronous Optical Network (SONET) rings by grooming low-speed traffic streams onto different high-speed wavelength channels. Since the cost of SONET add-drop multiplexers (SADM) at each node dominates the total cost of these networks, how to assign the wavelength, groom the traffic and bypass the traffic through the intermediate nodes has received a lot of attention from researchers recently. Moreover, the traffic pattern of the optical network changes from time to time. How to develop dynamic reconfiguration algorithms for traffic grooming is an important issue.

For dynamic traffic grooming (DTG) at the network operation stage, connection requests arrive and depart dynamically. As resources have already been deployed in the network and will remain unchanged for some time, the objective of a DTG algorithm is to maximize network throughput, or minimize the blocking probability of connection requests. To achieve this objective, the grooming algorithm must provision resource-efficient routes for both lightpaths and connections.

In this chapter, we describe, in detail, dynamic traffic grooming algorithms for two cases (best-fit and full-fit) for handling reconfigurable SONET over WDM networks. For each approach, an integer linear programming model and heuristic algorithms (based on the tabu search method) are given. The results demonstrate that the tabu search heuristic can yield better solutions but has a greater running time than the greedy algorithm for the best-fit case. For the full-fit case, the tabu search heuristic yields competitive results compared with an earlier simulated annealing based method and is more stable for the dynamic case. We also highlight related work on dynamic traffic grooming algorithms in other scenarios.

SONET rings have performed well as telecommunication backbone networks for a long time [1]. Using WDM technology, multiple rings can be

supported on a single fiber ring [2]. In this architecture, each wavelength independently carries a SONET ring. Each SONET ring can further support multiple low speed streams (e.g., an OC-48 SONET ring can support 4 OC-12 or 16 OC-3 streams at the same time). At each node a WDM Add/Drop Multiplexer (WADM) adds and drops or bypasses traffic on any wavelength. At each node, there are SONET add/drop multiplexers (SADM) on each wavelength to add/drop low speed streams. So the number of SADMs per node will increase linearly with the number of wavelengths that a single fiber ring can carry. The cost of SADMs will dominate the total cost of the optical network. But in fact, it is not necessary for each node to be equipped with SADMs on each wavelength. There is a need for an SADM on a wavelength at a node only if there is traffic terminating at this node on this wavelength. For example (see Fig. 12.1), if there is a traffic stream from node A to node B through node C on wavelength λ_1 , there should be an SADM on λ_1 at node A and node B. The traffic can bypass node C without add/drop capabilities for traffic streams on λ_1 . So node C need not be equipped with an SADM on λ_1 . The problem of combining different low speed traffic streams into high-speed traffic streams in such a way that the number of SADMs is minimized is called traffic grooming. Several studies have been done on traffic grooming [3, 4, 5].

Here is an example to show that traffic grooming can reduce the number of SADMs [2]. Consider a ring network with five nodes. Each wavelength could carry two traffic streams. The traffic pattern for this example is bidirectional uniform traffic. That is, there exists the same amount of traffic in both directions for each node pair. Below gives the traffic matrix of this example.

$$T = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{bmatrix}$$

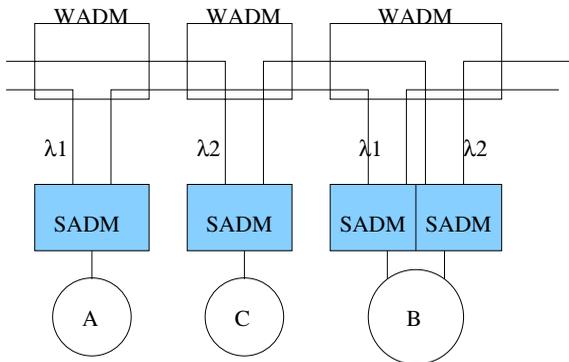


Fig. 12.1. SONET/WDM with bypass traffic at node C on λ_1

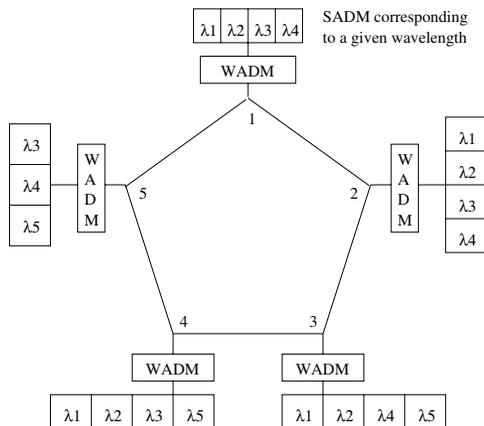


Fig. 12.2. SONET ring without traffic grooming

Figure 12.2 shows the SONET ring without traffic grooming. The traffic assignment is shown in Table 12.1. $1 \leftrightarrow 2$ in Tables 12.1 and 12.2 means that the traffic between the node 1 and 2 is groomed on the wavelength λ_1 . If there is a traffic stream from node 1 to node 3 assigned on a certain wavelength, we say that there is a virtual connection setting up from node 1 to node 3 on this wavelength. The total number of SADMs required is 19. Figure 12.3 shows the SONET ring with traffic grooming. The traffic assignment is shown in Table 12.2. The total number of SADMs is 15. In fact, if we do not know the traffic pattern and suppose there is an SADM at each node on each wavelength, the total number of SADMs is $5 \times 5 = 25$. Thus traffic grooming can reduce the number of SADMs greatly.

However, most algorithms assume the traffic matrix to be static; actually, the traffic pattern over SONET rings changes from time to time. In [6], dynamic traffic is described by a multiple set of the traffic matrices and a traffic grooming solution is proposed to meet the multiset instead of a single matrix. However, it is common that a change of traffic matrix happens after the configuration is established. In this paper, we consider the dynamic traffic grooming problem incorporating reconfiguration. That is, based on the current wavelength assignment, when the traffic pattern of the network

Table 12.1. Traffic assignment without traffic grooming

Wavelength	Traffic
λ_1	$(1 \leftrightarrow 2), (3 \leftrightarrow 4)$
λ_2	$(1 \leftrightarrow 3), (2 \leftrightarrow 4)$
λ_3	$(1 \leftrightarrow 4), (2 \leftrightarrow 5)$
λ_4	$(1 \leftrightarrow 5), (2 \leftrightarrow 3)$
λ_5	$(3 \leftrightarrow 5), (4 \leftrightarrow 5)$

Table 12.2. Traffic assignment with traffic grooming

Wavelength	Traffic
λ_1	$(1 \leftrightarrow 2), (1 \leftrightarrow 3)$
λ_2	$(1 \leftrightarrow 4), (1 \leftrightarrow 5)$
λ_3	$(2 \leftrightarrow 4), (2 \leftrightarrow 3)$
λ_4	$(4 \leftrightarrow 5), (2 \leftrightarrow 5)$
λ_5	$(3 \leftrightarrow 5), (3 \leftrightarrow 4)$

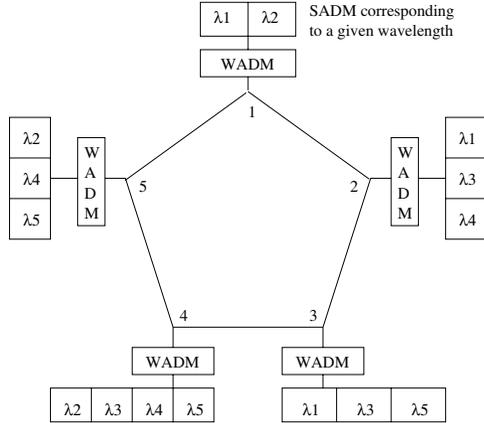


Fig. 12.3. SONET ring with traffic grooming

changes, we propose a dynamic traffic grooming algorithm to reconfigure the wavelength assignment according to the new traffic pattern without disrupting the old traffic assignment. Two cases, Best-fit and Full-fit, are studied. Two heuristic algorithms, Greedy and Tabu-search (TS-1) are presented for the best-fit approach. A two-phase algorithm based on tabu search (TS-2) is presented for the full-fit approach. The static traffic grooming problem, on which many studies have been done so far, is a special case of the *dynamic* problem (specially, the full-fit case). The simulation results illustrate that the tabu search algorithm can yield better solutions but takes more running time than greedy algorithms for the best-fit case. The tabu-search algorithm for the full-fit case is more stable compared to the earlier simulated annealing heuristic [3]. Some results for the static grooming problem are found to be better than earlier results in [3].

12.2 Problem Definition: Dynamic Traffic Grooming

In this section, integer linear programming models are proposed for dynamic traffic grooming. They are based on the static models proposed in [3]. There are two assumptions.

1. The old traffic matrix is known. By using a heuristic algorithm or Integer Linear Programming solver [3, 4], the current configuration of the network is obtained according to the old matrix. This corresponds to the initial assignment of traffic to wavelengths in the network.
2. The traffic matrix changes. The new matrix is different from the original matrix. The objective is to disrupt as few current connections as possible and fit the new traffic requests in.

The following are the notations that we will use later.

1. There are N nodes numbered $0, 1, 2, \dots, N - 1$ in the SONET ring.
2. W is the number of wavelengths in the original traffic grooming matrix.
3. The granularity of the traffic g is defined as:

$$g = \frac{\text{channel capacity}}{\text{base bandwidth rate}}$$

e.g., for a OC-48 channel carrying several OC-3 streams, the granularity $g = \frac{OC-48}{OC-3} = 16$. It is the number of circles¹ C that each wavelength can carry.

4. Original traffic matrix $T[n \times n]$. The traffic amount from node i to node j ($i, j = 1, \dots, N - 1, i \neq j$) on the ring is denoted by t_{ij} (entry of matrix T at row i and column j) and is always a multiple of the base bandwidth rate.
5. $T'[n \times n]$ is the new traffic matrix.
6. V_{ij}^{cw} represents the number of virtual connections for node pair (i, j) on wavelength w in circle c according to the original matrix. It is known in the reconfiguration problem.
7. $V'_{ij}{}^{cw}$ represents the number of new virtual connections from i to j on wavelength w of circle c according to the new traffic matrix.
8. V_{ij}^{cw+} represents the number of new virtual connections from i to j on new wavelength w^+ of circle c according to the new traffic matrix (applies to full-fit case only).
9. ADM_i^w represents the number of ADMs on node i on wavelength w for the original matrix.
10. $ADM'_i{}^w$ represents the number of additional ADMs on node i on wavelength w (applies to full-fit case only).
11. ADM_i^{w+} represents the number of ADMs on node i on new wavelength w^+ (applies to full-fit case only).

We assume that for the original traffic matrix, a solution has been found (e.g., using the algorithm in [3, 4]). So we know the number of wavelengths W and the virtual connections for each pair (i, j) . In addition, we also know the number of SADM and their positions.

Without special indication, unidirectional rings are assumed in the following problem descriptions. For bi-directional rings, constraints for both directions should be satisfied. Here are two cases to consider in the problem:

¹ The circle here refers to the circle built by the algorithm in [4].

12.2.1 Best-Fit Case

In best-fit case, we try to place as much new traffic as possible without increasing the number of SADMs.

Maximize : $\sum_i \sum_j \sum_c \sum_w V_{ij}^{\prime cw}$ (Objective function).

The objective is to maximize the traffic amount according to the new traffic demand. The following constraints are assumed.

1. Traffic constraint:

$$\sum_w \sum_c V_{ij}^{\prime cw} + \sum_w \sum_c V_{ij}^{cw} \leq t'_{ij}$$

The traffic constraint indicates that the total number of virtual connections from i to j should be less than the new demand traffic matrix of entry t'_{ij} .

2. Circle capacity constraint:

$$\sum_{(i,j)e \in (i,j)} (V_{ij}^{\prime cw} + V_{ij}^{cw}) \leq 1$$

The circle capacity constraint requires that no two connections can share a link on a circle.

3. Transmitter constraint:

$$\sum_c \sum_j V_{ij}^{\prime cw} + \sum_c \sum_j V_{ij}^{cw} \leq g \cdot ADM_i^w$$

The transmitter constraint requires that the total number of virtual connections starting at node i should be less than the transmitter capacity of SADM at this node on wavelength w .

4. Receiver constraint:

$$\sum_c \sum_i V_{ij}^{\prime cw} + \sum_c \sum_i V_{ij}^{cw} \leq g \cdot ADM_j^w$$

The receiver constraint requires that the total number of virtual connections terminating at node j should be less than the receiver capacity of SADMs at this node on wavelength w .

$V_{ij}^{cw'} \in \{0, 1, -1\}$. $ADM_j^w \in \{0, 1\}$. $V_{ij}^{cw'} = -1$ only if $V_{ij}^{cw} = 1$ and there is no connection between i and j on wavelength w of circle c for the new configuration any more.

12.2.2 Full-Fit Case

In full-fit case, we add the minimum number of SADM to satisfy all the new traffic.

Minimize : $\sum_i \sum_w ADM_i^w + \delta \sum_i \sum_{w^+} ADM_i^{w^+}$ (Objective function).

The objective is to fit all the traffic with the minimum number of SADMs added. δ in the objective function is the weight parameter representing the cost of adding more wavelengths. Because usually adding new wavelengths will cost more than adding more SADMs on existing wavelengths, δ is supposed to be no less than one. The following constraints are assumed.

1. Traffic constraint:

$$\sum_w \sum_c (V_{ij}^{cw} + V_{ij}^{\prime cw}) + \sum_{w^+} \sum_c V_{ij}^{cw^+} = t'_{ij}$$

The traffic constraint indicates that the total number of virtual connections from node i to node j should equal the traffic demand from node i to j .

2. Circle capacity constraint:

$$\sum_{(i,j) \in e \in (i,j)} (V_{ij}^{cw} + V'_{ij}{}^{cw}) \leq 1$$

$$\sum_{(i,j) \in e \in (i,j)} V_{ij}^{cw+} \leq 1$$

The circle capacity constraint requires that no two connections can share a single link on a circle.

3. Transmitter constraint:

$$\sum_c \sum_j (V_{ij}^{cw} + V'_{ij}{}^{cw}) \leq g \cdot (ADM_i^w + ADM_j'^w)$$

$$\sum_c \sum_j V_{ij}^{cw+} \leq g \cdot ADM_i^{w+}$$

The transmitter constraint requires that the total number of virtual connections should be less than the transmission capacity of the equipment at this node.

4. Receiver constraint:

$$\sum_c \sum_i (V_{ij}^{cw} + V'_{ij}{}^{cw}) \leq g \cdot (ADM_j^w + ADM_j'^w)$$

$$\sum_c \sum_i V_{ij}^{cw+} \leq g \cdot ADM_j^{w+}$$

The receiver constraint requires that the total number of virtual connections should be less than the receiving capacity of the equipment at this node.

$$V'_{ij}{}^{cw} \in \{0, 1, -1\}, V_{ij}^{cw+}, ADM_j'^w \text{ and } ADM_j^{w+} \in \{0, 1\}.$$

As we know, the integer linear programming problem is NP-complete [7]. The reconfiguration problem is described based on integer linear programming models. We expect this problem also to be intractable. In the next section, we propose heuristic approaches to solve this problem.

12.3 Heuristic Algorithms

The heuristic algorithms for dynamic grooming were developed for both the best-fit case and the full-fit case.

12.3.1 Best-Fit

The objective of the best-fit case is to include as much new traffic as possible using available capacity of the current configuration of the ring networks without increasing the number of SADMs. Two heuristic algorithms are proposed for best-fit case, greedy heuristic and Tabu search heuristic (TS-1).

Greedy Heuristic

In our greedy algorithm, the value of each entry of both the old matrix from which the current configuration was obtained and the new matrix using which

we will perform the reconfiguration is generated randomly in the range $[0, r]$. The value is uniformly distributed between 0 and r . We try to fit as much new traffic as possible without adding more SADMs. Here is a description of the algorithm.

1. Get grooming information for the original traffic matrix using an existing algorithm ([3, 4]).

For each circle, which we create for the current configuration, we should know whether the entire capacity between two nodes is occupied and whether there is an SADM at this node. We should also know on which wavelength this circle is groomed.

2. Find the difference traffic matrix.

Given the new traffic matrix, compute the difference between the old one and the new one, e.g., $D[i, j] = T'[i, j] - T[i, j]$. This is the matrix we try to groom in our algorithm with existing SADMs and traffic capacity. For some entries $D[i, j] = -m < 0$ (there exist some connections built for the old matrix that are not needed in the new matrix any more), remove the connections between (i, j) from m circles over at most m wavelengths.

3. Merge connections.

We want to keep the largest continuous gap between nodes. So, if there are two continuous connections over two circles, we merge them into a bigger one on the same circle.

4. Groom new traffic.

Start from the smallest hop.

From node 0 to node $N - 1$

From $s = 1$ to $s = N / *s$ is the number of hops*/

while ($D[i, i + s] \neq 0$)

 Try all the circles,

 for each node pair $(i, i + s)$ do

 if capacity is available *and*

 if there is SADM on both i and $i + s$

 {

 Groom one unit of $D[i, i + s]$ to this circle.

 Set capacity occupied.

$D[i, i + s] = D[i, i + s] - 1$.

 }

Endwhile

End

Tabu Search Heuristic (TS-1)

Tabu search is a meta-heuristic approach to solve hard optimization problems. The optimizing function is $f(x)$ subject to $x \in X$. The set X summarizes constraints on the vector of decision variables x . If x is the initial solution,

neighborhood $N(x)$ is a set obtained by going one step further from the solution x . Such a step is called a move. Each element in $N(x)$ is put into the candidate list. At the same time, a tabu list T is built to keep track of the solutions that have been visited before. Each element in the tabu list has a tabu tenure. After each move, the value of tabu tenure is decremented by one. Once the value becomes zero, the element is removed from the list. The tabu list prevents such solutions being revisited within cycles of length less than or equal to tabu tenure. A modified neighborhood $N^*(x)$ of current solution is $N(x)$ defined as $N(x) \setminus T$. Here $N^*(x)$ is a subset of $N(x)$. This kind of tabu search is short term TS. The best solution of $N^*(x)$ which is in the candidate list but not tabu (not in the tabu list) is chosen for the initial solution of next iteration. If after certain number of iterations (we call it the tabu limit) which is specified by the user, there is no improvement, the program stops. Otherwise, we continue the iteration and build a new candidate list. Here is the algorithm description of TS-1.

1. Initialization.

Compute the initial reconfiguration solution by using the greedy algorithm. Set the initial tabu tenure value, which should be defined before the program runs.

2. Build candidate list.

The neighborhood of the solution x , $N(x)$ is defined as:

$N(x) = \{x' \mid x' \text{ is a move by swapping part of the two circles of the solution } x \text{ and one part of the swap should have available capacity} \}$

Search all the circles. Compute $N(x)$ and put each element in $N(x)$ into the candidate list. Compute the additional new traffic amount that could be placed after such swap. Keep the maximum value of the increased traffic amount and its corresponding solution x .

3. Choose the best solution and continue.

For all the solutions in the candidate list, choose the move that is not tabu and can increase the traffic amount the most. The values of tabu tenure of all the elements in the tabu list are decremented by one. If some element's tenure becomes zero, remove it from the tabu list. Put the current move into the tabu list.

4. Terminate or continue with the current solution.

If after a certain number of iterations of moves, which is specified by the user, there is no improvement, the program stops. Otherwise, go back to build the candidate list for the new movement.

The tabu tenure for TS-1 is 48 and the tabu limit is 60.

12.3.2 Full-Fit Case

A two phase algorithm is developed for the full-fit case. The old traffic matrix and its solution are known. Recall that each entry of the old matrix and new

matrix is generated randomly ranging from $[0, r]$. The objective is to fit all new traffic requests. Adding more SADMs is allowed.

Algorithm Description

Here we present a two-phase algorithm for the full-fit case.

1. Use a best-fit algorithm (greedy or TS-1) to groom as much traffic as is allowed with existing SADMs.
2. If the capacity on circle c is available for the connection (i, j) , place SADMs at nodes i and j on the wavelength which circle c is groomed on and groom traffic onto this circle.
3. Use the tabu search algorithm (TS-2) to groom the remaining traffic onto the new wavelength. Place SADMs at the nodes whenever necessary.

Tabu Search Heuristic (TS-2)

The problem that was solved by TS-2 is the static traffic grooming problem, on which a lot of work has been done so far [3, 4, 5]. That is, given the remaining traffic matrix, groom that traffic onto wavelengths so that the number of SADMs is minimized. We observe that the static traffic grooming problem is a special case of the dynamic problem described in Section 12.2.2 when the old traffic matrix is empty and the new matrix is the traffic matrix that will be groomed. The following is the description of TS-2.

1. Initialization.
Use the algorithm in [4] to get an initial solution x . Set the initial value of the tabu tenure.
2. Build candidate list.
The neighborhood of the solution x , $N(x)$ is defined as:
 $N(x) = \{x' \mid x' \text{ is a move by swapping two circles from different wavelength of solution } x \text{ and the swapping will lead to no more or at most one more SADM on a wavelength}\}$.
Search all the circles. Compute $N(x)$ and place each element in $N(x)$ into the candidate list. Compute the number of SADMs that could be saved after this swap. Keep the maximum number of saved SADMs.
3. Choose the best solution and continue.
For all the solutions in the candidate list, choose the move that is not tabu and save the most SADMs. The values of tabu tenure of all the elements in the tabu list are decremented by one. If some element's tenure becomes zero, remove it from the tabu list. Place the current move into the tabu list.
4. Terminate or continue with the current solution.
If after a certain number of iterations of moves, which is specified by the user, there is no improvement, the program stops. Otherwise, go back to build the candidate list for the new movement.

The tabu limit is 170 and tabu tenure is in $\{15, 20, 24\}$.

12.4 Simulation Results and Analysis

In this section, we present the simulation results for both the best-fit case and the full-fit case.

12.4.1 Results for the Best-Fit Case

First, we define an upper bound on the number of new connections that can be groomed to the current configuration to evaluate the performance of the best-fit case algorithms. Then we give the result of both the greedy and the tabu search (TS-1) algorithms according to this upper bound. The running time comparison of both algorithms are also given.

Evaluation Factors

The results of reconfiguration algorithms depend not only on the algorithms themselves but also on the input matrix. The best-fit strategy strives to place as much traffic as it can, but there is no guarantee to fit all of the traffic in. Here we first develop an upper bound U on the number of connections that could possibly be groomed:

$$U = \sum_i \sum_j \sum_c \{V_{ij} \mid \text{there is enough capacity between } i \text{ and } j \text{ for new traffic } T'[i][j] \text{ and there is an SADM at both nodes } i \text{ and } j \text{ on circle } c\}.$$

The upper bound is computed by searching all the circles that are already built according to the old matrix to find available capacity to groom new connections. A connection can be established if there is capacity available between the terminal nodes and there are SADMs at the two nodes. The capacity of a connection is equal to the base bandwidth of one wavelength.

Although this upper bound is loose, because not all the connections available in the upper bound can be established at the same time, we can prove that no more connections can be built beyond this upper bound.

We define the load factor α by using this upper bound:

$$\alpha = \frac{\text{actual new connections established}}{\text{upper bound}} \times 100\%$$

This factor shows how much percentage of the new traffic could be groomed into the current configuration according to the upper bound.

Results of Greedy Heuristic

We randomly generate 18 matrix pairs for the old traffic matrices and the new matrices (the matrix pair is numbered as 1, 2, ..., 18 in Table 12.3 for a 5-node unidirectional ring with a granularity (g) of 3. The traffic amount generated for each node pair is evenly distributed between the range of 0 to 12 ($r = 12$). Table 12.3 provides the grooming results by using the evaluation factor α . From this table we observe that, when the number of nodes is not large and the granularity of the ring is small, the greedy algorithm can assign most of the new traffic that is possible to be groomed, according to the upper bound.

Table 12.3. Results for the best-fit greedy algorithms for 5 node ring

Matrix pair No.	1	2	3
Upper bound	15	24	24
Greedy connections	15	24	24
load factor α	100%	100%	100%
Matrix pair No.	4	5	6
Upper bound	23	20	13
Greedy connections	22	19	10
load factor α	96%	95%	77%
Matrix pair No.	7	8	9
Upper bound	10	19	8
Greedy connections	9	17	7
load factor α	90%	89%	88%
Matrix pair No.	10	11	12
Upper bound	23	12	18
Greedy connections	23	12	17
load factor α	100%	100%	94%
Matrix pair No.	13	14	15
Upper bound	22	18	33
Greedy connections	17	16	29
load factor α	82%	88%	88%
Matrix pair No.	16	17	18
Upper bound	18	16	32
Greedy connections	15	16	25
load factor α	83%	100%	78%

Results of Tabu Search Heuristic (TS-1)

Table 12.4 shows the results for the greedy algorithm and TS-1 under different numbers of nodes and different granularities for unidirectional rings ($r = 12$). For each entry, 20 matrix pairs (old matrix and new matrix) are randomly generated. The average value of α is computed. We observe that tabu search yields better results than the greedy algorithm for most cases, especially with a large number of nodes and large granularity. Table 12.5 shows the average number of new connections that can be groomed by the greedy algorithm and the tabu search algorithm using the same set of input data as Table 12.4. The tabu search heuristic gains 2% more connections than the greedy algorithm on the average.

Here we give a specific example to indicate that the tabu search heuristic can find a better solution than the greedy algorithm. In this example, there are 5 nodes in a unidirectional SONET ring. The granularity of the ring is 3. The old traffic matrix and the new matrix are generated as follows:

Table 12.4. Results for the best-fit algorithms. N is the number of nodes in a ring, α is the load factor. g is the granularity

Node number	method	g=3	g=4	g=12
N=5	greedy	92	89	79
(α in %)	tabu	92	90	80
N=6	greedy	90	89	78
(α in %)	tabu	91	90	80
N=8	greedy	90	84	76
(α in %)	tabu	91	85	77
N=12	greedy	86	83	70
(α in %)	tabu	87	84	72
N=16	greedy	85	82	67
(α in %)	tabu	86	84	68
N=20	greedy	84	80	63
(α in %)	tabu	86	82	69

$$T = \begin{bmatrix} 0 & 2 & 6 & 1 & 0 \\ 7 & 0 & 6 & 1 & 3 \\ 2 & 2 & 0 & 3 & 3 \\ 5 & 2 & 7 & 0 & 2 \\ 4 & 5 & 7 & 2 & 0 \end{bmatrix} \quad T' = \begin{bmatrix} 0 & 4 & 6 & 2 & 5 \\ 7 & 0 & 8 & 1 & 4 \\ 3 & 3 & 0 & 10 & 11 \\ 5 & 4 & 7 & 0 & 2 \\ 6 & 5 & 7 & 3 & 0 \end{bmatrix}$$

Figures 12.4 and 12.5 show the grooming results for the greedy algorithm and the tabu search algorithm, respectively. The following two matrices (R_g , R_t) are the remaining traffic matrices that cannot be groomed after employing those two algorithms. We observe from Fig. 12.5 that two connections of $t_{2,3}$ are groomed to wavelength 5 instead of wavelength 4 in the tabu search algorithm. Then two more connections of $t_{2,4}$ could be groomed to wavelength 4 which results in a gain of 2 more connections for the tabu search than the greedy heuristic.

Table 12.5. Absolute connections built by the greedy and the tabu algorithms

Node Number	method	g=3	g=4	g=12
N=5	greedy	16	17	22
	tabu	17	17	23
N=6	greedy	22	24	35
	tabu	22	25	36
N=8	greedy	39	40	54
	tabu	39	40	54
N=12	greedy	78	87	123
	tabu	79	88	126
N=16	greedy	128	142	193
	tabu	130	146	198
N=20	greedy	204	224	280
	tabu	209	229	290

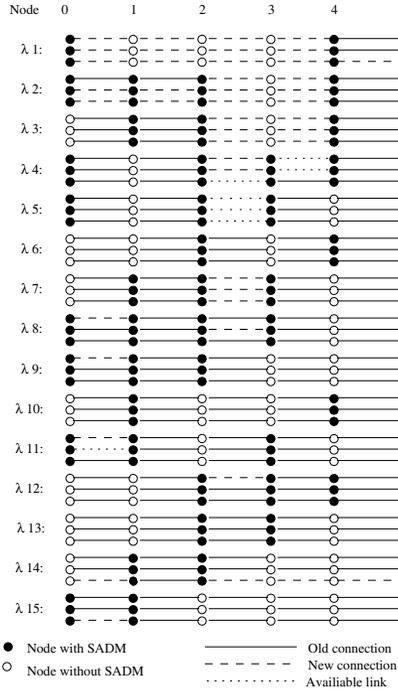


Fig. 12.4. An example of a reconfiguration problem using the greedy algorithm proposed in this work

$$R_g = \begin{bmatrix} 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 2 \\ 0 & 2 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \end{bmatrix} \quad R_t = \begin{bmatrix} 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \end{bmatrix}$$

Revenue Analysis and Running Time

While tabu search yields better results in most cases, it also takes more time to obtain the solution with the tabu heuristic than with the greedy algorithm. Figure 12.6 shows the running time of the algorithms for the unidirectional ring under different values of granularity and different numbers of nodes. It uses the same set of data that generated the result in Table 12.4. As we note from the figure, the running time increases greatly when the number of nodes in the ring increases. We compute the revenue [8] of the reconfiguration problem according to the number of connections that are built in total. Revenue is the income the network service provider earns by running the network. The following is the revenue formula for the reconfiguration problem.

C : the cost of using a single link (a single link is any node pair in $[i, i + 1]$)

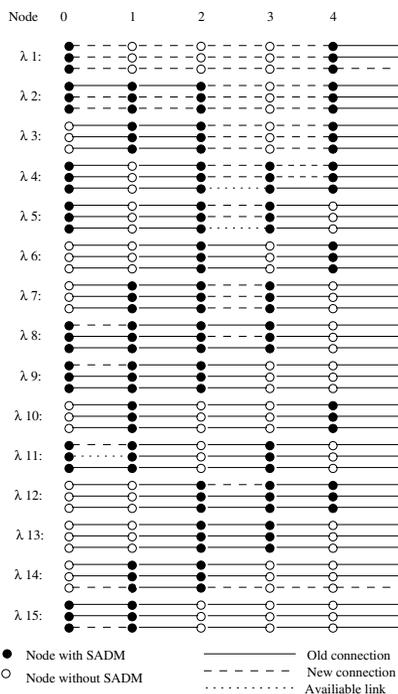


Fig. 12.5. An example of a reconfiguration problem using the tabu search heuristic (TS-1) proposed in this work

where $i \in 0..N - 1$).

$$\text{Revenue} = \sum_w \sum_c \sum_i \sum_j (V_{ij} + V'_{ij}) \times (j - i + N) \bmod N \times C \quad (12.1)$$

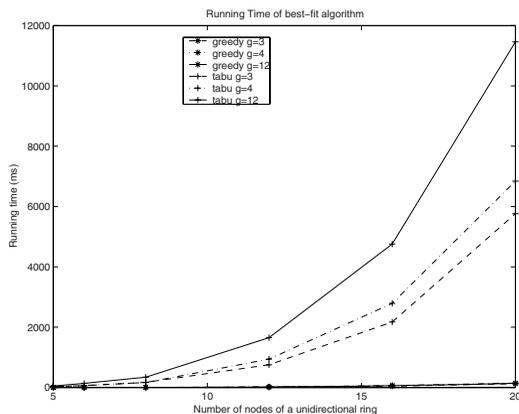


Fig. 12.6. Running time of the best-fit algorithm

From the formula, we observe that even one more connection can achieve quite a large percentage gain in the revenue if the cost of a single link is high. So, although the tabu search heuristic usually takes longer than the greedy algorithm to run, it is still worthwhile to employ tabu search when the single link cost is relatively high.

12.4.2 Results for the Full-Fit Case

We mentioned in Section 12.3.2 that the static traffic grooming problem is a special case of the dynamic problem (for the full-fit case). We run the algorithm (TS-2) for the static traffic grooming problem for a unidirectional ring under uniform traffic. The traffic amount of each entry in the matrix is one unit (equals base bandwidth). We compare the results with both the greedy algorithm [4] and the simulated annealing algorithm [3] in Table 12.6. For ILP results please refer to [3]. We observe that the tabu search algorithm obtains the same results as simulated annealing (SA) for most cases. One entry for the tabu search method is better than SA. For some cases, the number of SADMs used in tabu heuristic is a little bit more than in SA. The running time for each entry is less than 27 seconds on a 450 MHz UltraSPARC II processor based SUN Ultra-60 Workstation. In [3], the SA algorithm was run for 30 trials and best result was chosen. TS-2 does not depend on a statistical result and need not be run multiple times. We observe that it is never worse than the greedy algorithm. TS-2 is relatively stable, which is preferred in dynamic models. Because the traffic pattern changes from time to time, it is usually not feasible to run the program many times to obtain the best solution.

Table 12.6. Results for static traffic grooming

g	method	N=4	N=8	N=16	N=20
$g=3$	greedy	7	31	79	146
	SA	7	31	69	124
	TS-2	7	31	69	124
$g=4$	greedy	7	29	71	123
	SA	7	28	66	120
	TS-2	7	28	66	120
$g=16$	greedy	4	14	36	65
	SA	4	14	33	57
	TS-2	4	14	33	58
$g=48$	greedy	4	8	21	34
	SA	4	8	19	37
	TS-2	4	8	21	32
$g=64$	greedy	4	8	15	28
	SA	4	8	15	28
	TS-2	4	8	15	28

12.5 Conclusion

In this work, we introduced the dynamic traffic grooming model for reconfiguration problems. Two cases (best-fit and full-fit) are presented in an integer linear programming description. Since integer linear programming problems are NP-hard, we expect that dynamic traffic grooming problem is also intractable for both cases. For the best-fit case, two heuristic algorithms greedy and tabu search (TS-1) are proposed. An upper bound of the number of new virtual connections is developed to evaluate the performance of the algorithms. The results show that the tabu search algorithm (TS-1) proposed in this study will yield better solutions but takes more running time than the greedy algorithm for the best-fit case. For the full-fit case, a two-phase algorithm is developed. We observe that the static traffic grooming problem is a special case of the dynamic traffic grooming problem. The algorithms we proposed here (particularly the full-fit case, TS-2) can also solve the static grooming problem. Our algorithm is more stable than the simulated annealing algorithm proposed in previous work. Moreover, some of the solutions are better than those obtained in the earlier work on the static grooming problem.

As traffic grooming connects the optical layer and the electronic layer, a traffic grooming algorithm has to address not only the routing of traffic connections but also the routing and wavelength assignment (RWA) of wavelength channels. In other work, we proposed dynamic traffic grooming algorithms based on three different approaches. The first approach is an adaptive routing approach based on an auxiliary graph model [9]. The second approach is a fixed-alternate routing approach [10]. The third approach is a two-phase rerouting approach [11]. The common objective of these grooming algorithms is to maximize the network resource utilization or equivalently to minimize the connection blocking probability. We refer the interested reader to [12] and [13] and references therein for additional material on dynamic traffic grooming.

Acknowledgments The author thanks Ms. Shu Zhang and Dr. Wang Yao for their contributions to the the work presented here. This work was supported in part by the U.S. National Science Foundation (NSF).

References

1. Angela L. Chiu and Eytan H. Modiano (1998), "Reducing Electronic Multiplexing Costs in Unidirectional SONET/WDM Ring Networks via Efficient Traffic Grooming," *IEEE/IEICE Global Telecommunication Conference*, vol. 1, pp. 332–327, Sydney, Australia.
2. Eytan Modiano and Aradhana Narula-Tam (2000), "Mechanisms for providing optical bypass in WDM-based networks," *SPIE Optical Networks Magazine*.
3. Jian Wang, V. Rao Vemuri, Wonhong Cho and Biswanath Mukherjee (2000), "Improved Approaches for Cost-effective Traffic Grooming in WDM Ring networks: Non-uniform Traffic and Bidirectional Ring," *Proceedings ICC 2000*.

4. Xijun Zhang and Chunming Qiao (1998), "An Effective and Comprehensive Approach for Traffic Grooming and Wavelength Assignment in SONET/WDM Rings," *Proceedings of the SPIE-The International Society for Optical Engineering*, vol. 3531, pp. 221–232.
5. Ori Gerstel and Rajiv Ramaswami (1998), "Cost Effective Traffic Grooming in WDM rings", *Proceedings of IEEE INFOCOM '98*, vol. 1, pp. 69–77.
6. Randall Berry and Eytan Modiano (2000), "Reducing electronic multiplexing costs in SONET/WDM rings with dynamically changing traffic," *IEEE Journal on Selected Areas in Communications*, vol. 18, No. 10, pp. 1961–1971.
7. Alexander Schrijver (1986), *Theory of Linear and Integer Programming*, John Wiley & Sons, New York.
8. Murari Sridharan and Arun K. Somani (2000), "Revenue Maximization in Survivable WDM Networks," *Optical Networking and Communications, Proceeding of SPIE*, vol. 4233.
9. Wang Yao and Byrav Ramamurthy (2005), "A Novel Link Bundled Auxiliary Graph Model for Constrained Dynamic Traffic Grooming in Optical WDM Mesh Networks," *IEEE Journal of Selected Areas in Communications*, vol. 23, no. 8, pp. 1542–55.
10. Wang Yao and Byrav Ramamurthy (2004), "Dynamic traffic grooming using fixed-alternate routing in WDM mesh optical networks," in the *First Workshop on Traffic Grooming in WDM Networks, Co-located with BroadNets'2004*, San Jose, California.
11. Wang Yao and Byrav Ramamurthy (2004), "Rerouting schemes for dynamic traffic grooming in optical WDM mesh networks," in *IEEE Globecom 2004*, Dallas, Texas.
12. Wang Yao (2005), Traffic grooming in next-generation optical WDM mesh networks, Ph.D. dissertation, Department of Computer Science and Engineering, University of Nebraska-Lincoln.
13. Shu Huang and Rudra Dutta (2006), Dynamic traffic grooming: The changing role of traffic grooming, *IEEE Communications Surveys and Tutorials*, vol. 8, no. 4, pp. 32–50.
14. Peng-jun Wan, Liwu Liu and Ophir Frieder (1999), "Grooming of Arbitrary Traffic in SONET/WDM Rings," *IEEE/IEICE Global Telecommunication Conference*, vol. 1B, pp. 1012–1016.

Performance Models for Dynamic Traffic Grooming

Chunsheng Xin

13.1 Introduction

The traffic grooming problem originated from SONET and later extended into wavelength division multiplexing (WDM) optical networks. Traffic grooming in WDM optical networks is classified into *static* or *dynamic*, depending on whether the sub-wavelength client connections are given in advance, or randomly arrive/depart. Dynamic traffic grooming has been driven by several factors, e.g., the “bandwidth-on-demand” service needed by clients, and the demand to set up dynamic Label Switched Paths (LSPs) across optical networks, with the deployment of the Multi-Protocol Label Switching (MPLS) technology [1] in IP networks. While the major effort in static traffic grooming is to determine the lightpaths for aggregating the given client calls to optimize an objective, such as the number of accommodated calls [2] or the number of transponders [3], most studies on dynamic traffic grooming are to propose online algorithms to accommodate incoming sub-wavelength client connections on lightpaths that are set up on-demand in real time (e.g., see [4, 5, 6, 7, 8, 9, 10]). It is important to develop performance models for dynamic traffic grooming to evaluate the performance of grooming algorithms as well as help to study other problems in dynamic traffic grooming. A major performance metric of dynamic traffic grooming is the blocking probability of sub-wavelength connections. As such, we only discuss the analysis of blocking probability.

There is a large number of analytical studies on WDM optical networks with wavelength-granularity traffic demands, e.g., see [11, 12, 13, 14, 15, 16] and references therein. Nevertheless, these studies assumed wavelength-granularity traffic demands, and did not address sub-wavelength traffic demands. The authors in [17, 18] analyzed *WDM grooming* in time-space switched optical networks, where optical nodes are capable of sub-wavelength time-slot switching, and sub-wavelength connections can be directly set up over the time-slots in wavelength channels. In traffic grooming, optical nodes switch traffic only at the wavelength granularity, and

sub-wavelength connections have to be carried on lightpaths, e.g., see [2, 3, 4, 5, 6, 7, 8, 9, 10, 19, 20, 21]. We focus on this scenario and discuss appropriate performance models. We will utilize an approximation technique known as *reduced load approximation*, which was developed in [22, 23, 24] and has been widely adopted for analyzing general networks and optical networks.

13.2 Characterization of Dynamic Traffic Grooming

We use *calls* to refer to dynamic sub-wavelength connections from clients. The physical topology in dynamic traffic grooming includes the optical network, and client nodes connected to border optical nodes, as illustrated in Fig. 13.1. Lightpaths can be set up (and torn down) in real time between a *pair of client nodes* (node-pair). Upon receiving an incoming call request, a client node selects a sequence of lightpaths to carry this call to the destination. In this lightpath sequence, some lightpaths may not exist yet and need to be set up in real time. The grooming algorithm also tears down a lightpath if all carried calls are completed. The dynamic grooming algorithms can be generally classified into *single-hop* or *multi-hop* algorithms, depending on whether the lightpath sequence can include only one lightpath or multiple lightpaths. The single-hop grooming algorithm aggregates a call on a single lightpath to eliminate intermediate electronic processing (e.g., see [21, 25]). On the other hand, a multi-hop grooming algorithm may aggregate a call on several lightpaths, to reduce blocking probability. As a trade-off, the call needs to go through intermediate client nodes to be processed and forwarded in the electronic layer.

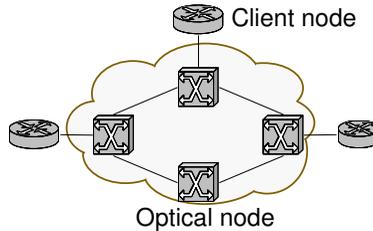


Fig. 13.1. Physical topology

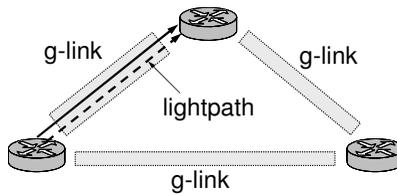


Fig. 13.2. g-link topology

The lightpaths set up in the optical network connect client nodes into a logical topology. Dynamic traffic grooming can be seen as client call routing on a dynamic logical topology, where logical links (i.e., lightpaths between node-pairs) are dynamically set up or torn down to increase or decrease capacity between a node-pair. It is difficult to analyze call blocking probabilities on a dynamic logical topology. One solution is to introduce the concept of *grooming link* (g-link) and *g-link topology* to transform a dynamic logical topology to an equivalent “static” topology. A g-link is conceived as a virtual and static “conduit” used to enclose the dynamically established lightpaths between a node-pair. The g-links between all node-pairs form a fully connected and static *g-link topology*, as illustrated in Fig. 13.2. The g-link topology together with the information on the number of lightpaths in each g-link uniquely defines a logical topology.

The single and multi-hop grooming algorithms can easily be described on the g-link topology. A single-hop grooming algorithm essentially routes a call on a single-hop path in the g-link topology, while a multi-hop grooming algorithm routes a call on a path with one or more hops.

13.3 Challenges and Solutions

To avoid confusion, a wavelength channel in a particular fiber is referred to as a *wavelink*, and a wavelink route that corresponds to a possible wavelength assignment on a fiber route is referred to as a *waveroute*. Figure 13.3 illustrates two waveroutes corresponding to two wavelength assignments on a fiber route, assuming two wavelengths per fiber and no wavelength conversion. The entire set of waveroutes (on all fiber routes) used to establish lightpaths in a g-link is referred to as the *support* of this g-link. The call loss probability between a node-pair can be calculated based on the blocking probabilities of the routing paths on the g-link topology, which are dependent on the g-link blocking probabilities. However, the number of lightpaths and the capacity of a g-link are dynamic, and it is difficult to directly calculate the g-link blocking probabilities. This problem can be solved by modeling the two operations for grooming calls on a g-link—routing calls to lightpaths and dynamically establishing lightpaths on wavelength channels for this g-link—by a single operation of routing calls directly on wavelength channels. Then the call blocking probability on a g-link becomes the joint blocking probability of the entire

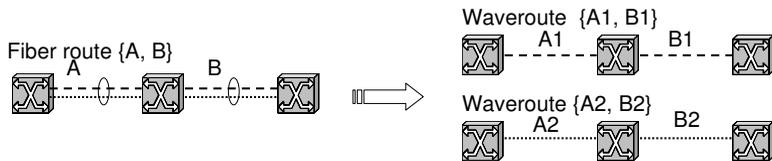


Fig. 13.3. Relationship between fiber route and waveroute

set of waveroutes supporting this g-link. However, for the single operation to have the same effect as the two operations, it has to meet two constraints.

1. Routing a call on wavelength channels must follow the routing and wavelength assignment (RWA) algorithm used to set up lightpaths, because calls are actually carried on lightpaths. Specifically, calls arriving at a g-link are routed to waveroutes that are on the fiber routes selected by the RWA algorithm to set up lightpaths in this g-link, and the waveroute selection follows the RWA scheme.
2. Calls carried on a wavelink must be on the same waveroute, to ensure that they are actually contained in a lightpath, because the traffic carried on a wavelink is actually the traffic on a lightpath that occupies the wavelink.

The second constraint can be satisfied by using a *restricted sharing* admission policy for the wavelink, which admits an incoming call if and only if the wavelink is in the empty state, or the waveroute of the incoming call is the same as the one of the carried calls and there is enough residual bandwidth.

13.4 A General Analysis Framework

Theoretically a Markov process might be used for the analysis of dynamic traffic grooming, but is far too complex. In practice, approximation techniques need to be employed. A widely used technique in network analysis called the *reduced load approximation* (see [22, 23, 24, 26]) can be adopted to analyze dynamic traffic grooming.

The objective of dynamic traffic grooming analysis is to estimate the call blocking probability for given offered loads to each node-pair. Based on the discussion in Section 13.3, we illustrate dynamic traffic grooming in Fig. 13.4. Logically, client calls are carried on the paths composed of g-links. On the other hand, client calls carried on a g-link are physically carried on waveroutes supporting this g-link, and thus carried on wavelinks. Using this hierarchy, the blocking analysis for dynamic traffic grooming is composed of three

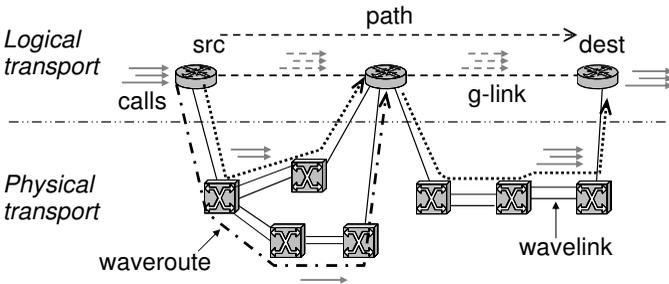


Fig. 13.4. Hierarchy in dynamic traffic grooming

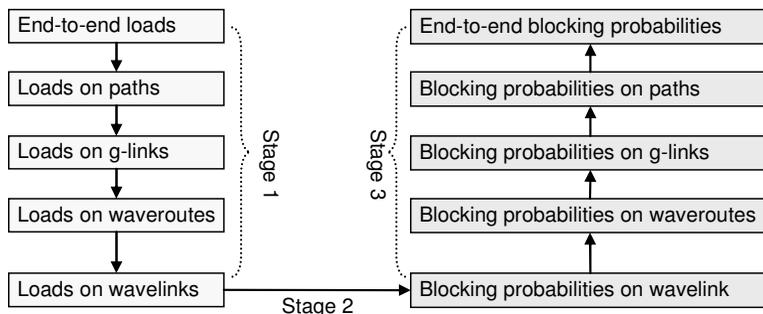


Fig. 13.5. Calculate call loss probabilities from end-to-end offered loads

stages: (1) derive wavelink offered loads from the given offered loads between node-pairs; (2) compute wavelink blocking probabilities from wavelink offered loads using a wavelink blocking model; and (3) calculate the end-to-end call loss probabilities from wavelink blocking probabilities. Figure 13.5 illustrates this procedure, where the blocking probability of a g-link is equal to the joint blocking probability of all waveroutes supporting this g-link, and the end-to-end call loss probability between a node-pair is equal to the joint blocking probability of all g-link paths used to route calls between this node-pair.

In the first stage, the end-to-end load of a node-pair is partitioned and offered to the g-link paths between this node-pair, e.g., through a sequential overflowing. The traffic load carried on a g-link is the summation of the carried loads on all paths traversing this g-link. The traffic arriving at a g-link is partitioned and offered to the waveroutes supporting this g-link, e.g., through a sequential overflowing or load sharing (each waveroute gets the same amount of traffic). Then the traffic load traversing a wavelink is simply the traffic on the waveroutes traversing this wavelink. However, this stage is a little tricky because when computing these traffic loads, the corresponding blocking probabilities are required. For instance, to calculate the traffic load carried on a specific path, we need to know both the traffic load offered to this path and the blocking probability of this path, as illustrated in Fig. 13.6. Nevertheless, our objective is to compute the blocking probabilities from these traffic loads. In other words, the blocking probabilities shown in Fig. 13.5 not only depend on the traffic loads, but also depend on themselves, which results in a set of nonlinear equations called *fixed-point equation* in the literature (e.g., see [22, 23, 24]). The fixed-point equation can be solved by a nonlinear solver and the end-to-end call loss probabilities can then be obtained.

The wavelink blocking model plays a fundamental role to obtain blocking probabilities from traffic loads. The next section introduces such a model. After this, a basic blocking model is given to illustrate the entire procedure in the analysis of traffic grooming.

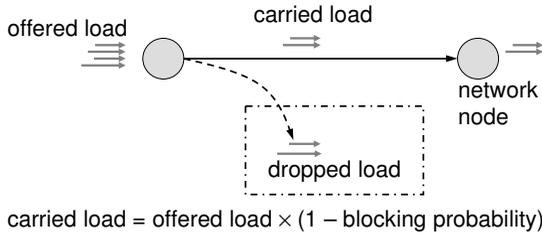


Fig. 13.6. Calculation of carried load

13.5 Wavelink Blocking Model

In traffic grooming, the calls are likely heterogeneous (with different data rates and possibly other different requirements), and hence the calls should be assumed multi-service calls, where a *service* primarily represents a data rate, but can indicate other additional requirements. The multi-service calls arriving at a wavelink are assumed following a Poisson process and having exponentially distributed holding times. In the restricted sharing admission policy¹ that is used for the wavelink, the waveroute traversed by a call is a required information for a wavelink to accept or block this call. Thus we must differentiate calls by their waveroutes. As such, we define the *class* of a call as the combination of the waveroute it traverses and its service type.

Let N denote the number of waveroutes that traverse the wavelink and $\mathcal{R}_1, \dots, \mathcal{R}_N$ denote these waveroutes. Let \mathcal{K} denote the set of call services and b_k denote the data rate of the service- k calls. Let C be the wavelink capacity. Let ρ_{nk} denote the offered load of class- $\langle n, k \rangle$ calls, i.e., the service- k calls on waveroute \mathcal{R}_n , arriving at the wavelink.

Theorem 1. *Let $q(n, j)$ ($1 \leq n \leq N, 1 \leq j \leq C$) be the probability that the wavelink is occupied by calls on waveroute \mathcal{R}_n and exactly j units of bandwidth are used in the wavelink. Let $q(0)$ be the probability that the wavelink is empty (no calls).*

Then $q(n, j)$ satisfies the relation

$$j \cdot q(n, j) = \sum_{k \in \mathcal{K}} b_k \cdot \rho_{nk}(j - b_k) \cdot q(n, j - b_k), \tag{13.1}$$

where $q(n, j) = 0$ for $j < 0$, $q(n, 0) = q(0)$ and

$$q(0) + \sum_{n,j} q(n, j) = 1. \tag{13.2}$$

¹ See Section 13.3.

Denote the wavelink blocking probability for class- $\langle n, k \rangle$ calls as B_{nk} . Then

$$B_{nk} = 1 - q(0) - \sum_{j=1}^{C-b_k} q(n, j), \quad (13.3)$$

Proof. See [21].

13.6 A General Blocking Model for Dynamic Traffic Grooming

Now we present a general blocking model. It can be used for general grooming algorithms that use fixed or alternate routing. The following assumptions are made:

- The optical network has no wavelength conversion.
- Link-disjoint alternate routing for both lightpaths (in the physical topology) and client calls (in the g-link topology).
- Sequential wavelength assignment, e.g., first-fit.
- Independent link blocking for wavelinks and g-links.
- A blocked lightpath/client call after trying all fiber routes/g-link paths is lost.

Based on the RWA algorithm, the waveroutes supporting each g-link are pre-computed for the analysis model. Let J denote the number of alternate g-link paths between a node-pair and U denote the number of waveroutes supporting a g-link. Let W denote the number of wavelinks in the physical topology. We use the following notations in the analysis model.

a^{sk}	Load of service- k calls offered to node-pair s .
\mathcal{L}_{sk}	Loss probability of service- k calls offered to node-pair s .
\mathcal{L}_{ave}	Average loss probability of calls (of all services) offered to the network (all node-pairs).
\mathcal{P}_{sj}	The j th path between node-pair s in the g-link topology.
\mathcal{B}^{tk}	Blocking probability of g-link t for service- k calls.
G_k^{sj}	Blocking probability of path \mathcal{P}_{sj} for service- k calls.
θ_k^{sj}	Load of service- k calls carried on path \mathcal{P}_{sj} .
C^{tk}	Load of service- k calls carried on g-link t .
A^{tk}	Load of service- k calls offered to g-link t .
\mathcal{R}_{tu}	The u th waveroute supporting g-link t .
L_k^{tu}	Blocking probability of waveroute \mathcal{R}_{tu} for service- k calls.
E_k^{tu}	Load of service- k calls carried on waveroute \mathcal{R}_{tu} (i.e., class- $\langle t, u, k \rangle$ calls).
B_w^{tuk}	Blocking probability of class- $\langle t, u, k \rangle$ calls on wavelink w .
D_w^{tuk}	Load of class- $\langle t, u, k \rangle$ calls offered to wavelink w .

13.6.1 Calculation of Traffic Loads in Stage 1

We first discuss the procedure to obtain the traffic loads on each quantity in the hierarchy discussed in Section 13.4. The calculation of the traffic loads needs the blocking probabilities. Although the actual values of these quantities are not available at this moment, they can be treated as unknowns in functions in the derivation of traffic loads.

By the alternate routing assumption, a call of node-pair s is sequentially overflowed among paths between this node-pair, $\mathcal{P}_{s,1}, \dots, \mathcal{P}_{s,J_s}$. The carried load of a path is then equal to the offered load to this path times the accepting probability, illustrated in Fig. 13.6. Thus the path \mathcal{P}_{s_j} carried load of service- k calls ($\theta_k^{s_j}$) is calculated as

$$\theta_k^{s_j} = a^{sk} \cdot \prod_{i=1}^{j-1} G_k^{s,i} \cdot (1 - G_k^{s,j}), \quad (13.4)$$

where $\prod_{i=1}^{j-1} G_k^{s,i}$ is the joint blocking probability of the first $j-1$ paths $\mathcal{P}_{s,1}, \dots, \mathcal{P}_{s,j-1}$, and thus $a^{sk} \cdot \prod_{i=1}^{j-1} G_k^{s,i}$ is the offered load to path $\mathcal{P}_{s,j}$.

The carried load on a g-link is the summation of the carried loads of all paths traversing this g-link. Let $\mathcal{Q}^t = \{(s, j)\}$ denote the indices of the paths that traverse g-link t . Then the g-link t carried load of service- k calls can be calculated as follows

$$C^{tk} = \sum_{(s,j) \in \mathcal{Q}^t} \theta_k^{s_j}. \quad (13.5)$$

The g-link t offered load (A^{tk}) can then be computed from the g-link t carried load as below (see Fig. 13.6).

$$A^{tk} = \frac{C^{tk}}{1 - \mathcal{B}^{tk}} \quad (13.6)$$

By the alternate routing and sequential wavelength assignment, a service- k call arriving at g-link t sequentially tries the waveroutes supporting g-link t , $\mathcal{R}_{t,1}, \dots, \mathcal{R}_{t,U}$. With the assumption of link-disjoint routing and no wavelength conversion in the optical network, the waveroutes $\mathcal{R}_{t,1}, \dots, \mathcal{R}_{t,U}$ are wavelink-disjoint and thus their blocking probabilities are independent. The carried load of service- k calls on waveroute \mathcal{R}_{t_u} can be computed as follows, similar to the calculation in (13.4).

$$E_k^{t_u} = A^{tk} \cdot \prod_{i=1}^{u-1} L_k^{t,i} \cdot (1 - L_k^{t,u}), \quad (13.7)$$

where $\prod_{i=1}^{u-1} L_k^{t,i}$ is the joint blocking probability of the first $u-1$ waveroutes $\mathcal{R}_{t,1}, \dots, \mathcal{R}_{t,u-1}$.

The carried load of class- $\langle t, u, k \rangle$ calls on a wavelink is the carried load of service- k calls on waveroute \mathcal{R}_{tu} , i.e., E_k^{tu} , if \mathcal{R}_{tu} traverses this wavelink. Thus the offered load of class- $\langle t, u, k \rangle$ calls to wavelink w (D_w^{tuk}) can be calculated as follows (see Fig. 13.6):

$$D_w^{tuk} = \begin{cases} \frac{E_k^{tu}}{1-B_w^{tuk}} & \text{if } w \in \mathcal{R}_{tu} \\ 0 & \text{otherwise} \end{cases}. \quad (13.8)$$

13.6.2 Calculation of Blocking Probabilities in Stage 3

After obtaining wavelink offered loads D_w^{suk} for all s, u, k , and w , wavelinks blocking probabilities B_w^{suk} can be computed using the blocking model in Section 13.5. The blocking probabilities in the third stage can then be computed.

By the independent link blocking assumption, the blocking probability of waveroutes R_{tu} for service- k calls is

$$L_k^{tu} = 1 - \prod_{w \in \mathcal{R}_{tu}} (1 - B_w^{tuk}). \quad (13.9)$$

As discussed earlier, the g-link blocking probability is equal to the joint blocking probability of all waveroutes supporting this g-link. Thus the blocking probability of g-link t (\mathcal{B}^{tk}) is as follows:

$$\mathcal{B}^{tk} = \prod_{i=1}^U L_k^{ti}. \quad (13.10)$$

By the independent link blocking, the blocking probability of path \mathcal{P}_{sj} for service- k calls is

$$G_k^{sj} = 1 - \prod_{t \in \mathcal{P}_{sj}} (1 - \mathcal{B}^{tk}). \quad (13.11)$$

By the alternate routing assumption, a call of node-pair s is sequentially overflowed among paths $\mathcal{P}_{s,1}, \dots, \mathcal{P}_{s,J}$. Thus the end-to-end call loss probability of node-pair s is

$$\mathcal{L}_{sk} = \prod_{i=1}^J G_k^{s,i}. \quad (13.12)$$

The average end-to-end call loss probability among all node-pairs is obtained by summing up dropped and offered loads on all call services and all node-pairs. Thus we have

$$\mathcal{L}_{ave} = \sum_{s,k} a^{sk} \mathcal{L}_{sk} / \sum_{s,k} a^{sk}. \quad (13.13)$$

13.6.3 Nonlinear Dependence of Blocking Probabilities

The above computation from the offered loads of node-pairs to call loss probabilities between node-pairs is not a linear procedure, since the calculation of traffic loads in Stage 1 are dependent on the blocking probabilities to be computed in Stage 3. Since all blocking probabilities in Stage 3 are calculated based on wavelink blocking probabilities, and wavelink offered loads are the final result in Stage 1, we represent the dependence of traffic loads on blocking probabilities in a function form $\mathbf{D}(\mathbf{B})$, where \mathbf{D} and \mathbf{B} represent the matrix forms of D_w^{tuk} and B_w^{tuk} . Let $\varphi(\bullet)$ denote the wavelink blocking model as a function form. Then the nonlinear dependence results in the following equation

$$\mathbf{B} = \varphi(\mathbf{D}(\mathbf{B})) = \varphi'(\mathbf{B}),$$

which is the *fixed-point equation* mentioned in Section 13.4. This equation can be solved by a nonlinear solver, e.g., through the repeated substitution illustrated in Algorithm 1. After it converges, the end-to-end call loss probabilities are obtained by (13.12)–(13.13).

We give some numerical results of the above model. A 14-node sample network illustrated in Fig. 13.7 is used. Each optical node is attached to one client node. At each client node, client calls are generated based on the Poisson arrival and exponential holding time assumptions. The wavelink capacity is assumed 100. The offered load between a node-pair is uniformly distributed

Algorithm 1 Compute call loss probabilities

1. Assign arbitrary numbers in $(0, 1)$ to B_w^{tuk} for all t, u, k, w
 2. Compute L_k^{tu} for all t, u, k by (13.9)
 3. Compute \mathcal{B}^{tk} for all k, t by (13.10)
 4. Compute G_k^{sj} for all s, j, k by (13.11)
 5. Compute θ_k^{sj} for all s, j, k by (13.4), with the given a^{sk}
 6. Compute \mathcal{C}^{tk} , and A^{tk} for all t, k by (13.5) and (13.6)
 7. Compute E_k^{tu} for all t, u, k by (13.7)
 8. Compute D_w^{tuk} for all t, u, k, w by (13.8)
 9. Compute B_w^{tuk} for all t, u, k, w by (13.1)–(13.3). If B_w^{tuk} have converged, terminate. Otherwise go to step 2
-

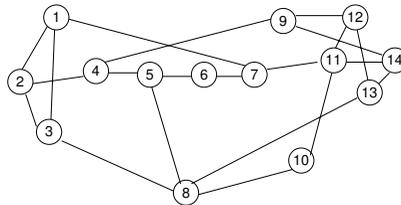


Fig. 13.7. A 14-node sample network

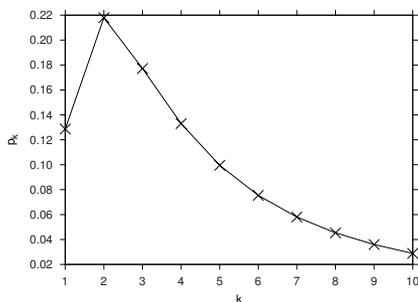


Fig. 13.8. Log normal distribution

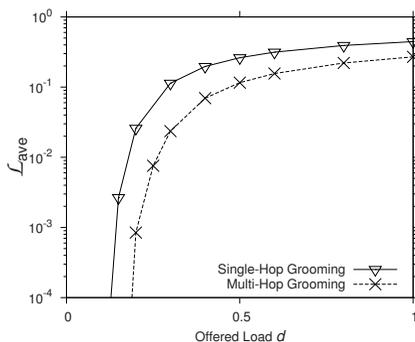


Fig. 13.9. Blocking probability

in the range $[1, 1 + d]$, where d is a simulation parameter. The number of call services is assumed 10, and the data rate of service- k ($1 \leq k \leq 10$) calls is assumed $10k$. The probability of an incoming call being a service- k call follows the log-normal distribution as illustrated in Fig. 13.8. The shortest path routing and first-fit wavelength assignment are assumed in the optical network. The number of wavelengths is assumed 8. The single-hop traffic grooming algorithm and a multi-hop grooming algorithm are used, where the latter utilizes the single-hop path and up to 4 multi-hop paths computed through setting the g-link distance as the length of the waveroute supporting the g-link. Figure 13.9 plots the computed blocking probabilities.

13.7 Extensions of the Basic Blocking Model

The blocking model in Section 13.6 can be extended to address other features and requirements in dynamic traffic grooming. We describe four examples to extend the model to address wavelength conversion, different traffic grooming policy, arbitrary alternate routing, and random traffic grooming.

13.7.1 Wavelength Conversion in Dynamic Traffic Grooming

The optical network might have wavelength conversion. With wavelength conversion, waveroutes corresponding to a fiber route overlap with each other to create *repeated wavelinks* (see Section 13.7.3), since the waveroutes are the possible wavelength assignments on this fiber route. Furthermore, the number of waveroutes corresponding to a fiber route increases exponentially as a function of the number of conversion nodes on the route. Together, they make the computation of the blocking probabilities time-consuming. To improve the efficiency, an approximation technique was proposed in [27]. A fiber route is decomposed into a sequence of *segments*,² a partial route between two consecutive conversion nodes or from the source/destination to the closest conversion node on a fiber route. As shown in Fig. 13.10, a segment is composed of multiple link-disjoint waveroutes since there is no conversion node inside a segment.

The calls offered to a fiber route arrive at each segment of the route, and are then offered to the waveroutes on each segment. The sub-wavelength calls eventually arrive at the wavelinks of these waveroutes. With this traffic grooming modeling, the blocking probability of a g-link can be calculated from the blocking probabilities of the fiber routes used to set up lightpaths for these g-links. On the other hand, the blocking probability of a fiber route can be calculated from the blocking probabilities of the segments on the route, which can be obtained from the blocking probabilities of the waveroutes on these segments. Therefore, the call loss probability between a node-pair can be computed similarly as in Section 13.6, through a hierarchy of wavelink-waveroute-segment-g-link-path. The number of waveroutes on the segments of a fiber route increases only linearly as a function of the number of conversion nodes in the fiber route. Furthermore, they are wavelink-disjoint, and also shorter than the end-to-end waveroutes on the fiber route. Therefore, the segment blocking probability can be efficiently computed.

We briefly discuss the modification to the basic blocking model in Section 13.6 for this scenario. First we redefine the following notations to make them correspond to quantities of fiber routes and then describe how to obtain them. Other notations and their computations are the same.

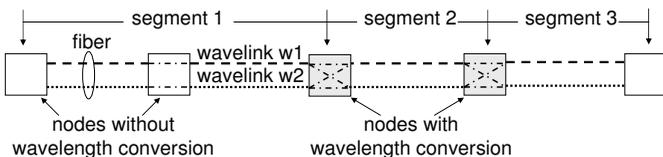


Fig. 13.10. Decomposing a fiber route into segments

² In contrast, the approach taken in Section 13.6 decomposes a fiber route into a set of end-to-end waveroutes, and the fiber route blocking probability is computed as the joint blocking probability of all waveroutes on the fiber route.

U	Number of fiber routes to set up lightpaths for a g-link.
\mathcal{R}_{tu}	The u th fiber route used to set up lightpaths on g-link t , consisting of segments.
B_w^{tuk}	Blocking probability of wavelink w for class- $\langle t, u, k \rangle$ calls (service- k calls on fiber route \mathcal{R}_{tu}).
$\mathbb{B}_{\mathbb{R}}^{tuk}$	Blocking probability of segment \mathbb{R} for class- $\langle t, u, k \rangle$ calls.
L_k^{tu}	Blocking probability of fiber route \mathcal{R}_{tu} for service- k calls.
E_k^{tu}	Carried load of service- k calls on fiber route \mathcal{R}_{tu} .
$H_{\mathbb{R}}^{tuk}$	Offered load of class- $\langle t, u, k \rangle$ calls arriving at segment \mathbb{R} .

The offered load to a g-link is offered to the fiber routes between this g-link. Furthermore, by the link-disjoint alternate routing assumption, the carried load of service- k calls on route \mathcal{R}_{tu} (i.e., class- $\langle t, u, k \rangle$ calls) is

$$E_k^{tu} = A^{tk} \cdot \prod_{i=1}^{u-1} L_k^{ti} \cdot (1 - L_k^{tu}), \quad (13.14)$$

where A^{tk} is the g-link offered load (see Section 13.6) and $\prod_{i=1}^{u-1} L_k^{ti}$ is the joint blocking probability of first $u - 1$ fiber routes.

Based on the route carried load, the offered load of class- $\langle t, u, k \rangle$ calls arriving at segment \mathbb{R} can then be calculated as

$$H_{\mathbb{R}}^{tuk} = \begin{cases} \frac{E_k^{tu}}{1 - \mathbb{B}_{\mathbb{R}}^{tuk}} & \text{if } \mathbb{R} \in \mathcal{R}_{tu} \\ 0 & \text{otherwise} \end{cases}. \quad (13.15)$$

For $\mathbb{R} \in \mathcal{R}_{tu}$, denote $|\mathbb{R}|$ as the number of waveroutes in segment \mathbb{R} (note that \mathbb{R} is composed of link-disjoint waveroutes). Let $r_1, \dots, r_{|\mathbb{R}|}$ be the waveroutes of \mathbb{R} . A call arriving at a segment tries its waveroutes sequentially, and thus the carried load of service- k calls on waveroute r_i , denoted as V_i , is

$$V_i = \prod_{\ell \in r_i} (1 - B_{\ell}^{tuk}) \cdot H_{\mathbb{R}}^{tuk} \cdot \prod_{j=1}^{i-1} \left(1 - \prod_{\ell \in r_j} (1 - B_{\ell}^{tuk}) \right),$$

where the first term is the call accepting probability on waveroute r_i , and the last term is the joint blocking probability of the first $i - 1$ waveroutes r_1, \dots, r_{i-1} on segment \mathbb{R} .

The carried load of class- $\langle t, u, k \rangle$ calls on every wavelink $w \in r_i$ is the same as the carried load of service- k calls on waveroute r_i . Thus the offered load of class- $\langle t, u, k \rangle$ calls arriving at the wavelink $w \in r_i$ ($1 \leq i \leq |\mathbb{R}|$) can be calculated as

$$D_w^{tuk} = \frac{V_i}{1 - B_w^{tuk}}.$$

After obtaining the wavelink offered loads D_w^{tuk} for all t, u, k, w , wavelink blocking probabilities B_w^{tuk} can be calculated using the wavelink blocking

model, and then other blocking probabilities can be computed based on them as follows.

By the sequential wavelength assignment assumption, a call arriving at a segment sequentially tries its waveroutes. Let $r \in \mathbb{R}$ denote a waveroute in segment \mathbb{R} . Then the call blocking probability of \mathbb{R} for class- $\langle t, u, k \rangle$ calls is calculated by the independent link blocking assumption as

$$\mathbb{D}_{\mathbb{R}}^{tuk} = \prod_{r \in \mathbb{R}} \left(1 - \prod_{w \in r} (1 - B_w^{tuk}) \right). \quad (13.16)$$

The blocking probabilities of segments on a fiber route are independent, and thus the blocking probability of fiber route \mathcal{R}_{tu} for service- k calls is

$$L_k^{tu} = 1 - \prod_{\mathbb{R} \in \mathcal{R}_{tu}} (1 - \mathbb{D}_{\mathbb{R}}^{tuk}).$$

As discussed earlier, the blocking probability of a g-link is now the joint blocking probability of the fiber routes that are used to set up lightpaths for this g-link. Thus the blocking probability of g-link t for service- k calls is

$$\mathcal{B}^{tk} = \prod_{i=1}^U L_k^{ti}. \quad (13.17)$$

Other blocking probabilities are calculated as in Section 13.6. Similarly a fixed-point equation is formed and after solving this equation, the call loss probabilities between node-pairs (\mathcal{L}_{sk}) can be obtained.

13.7.2 Traffic Grooming Policy

When routing a call onto a multi-hop g-link path, some traffic grooming algorithms may want to use only existing lightpaths on these paths, so that the source client node does not need to request to set up a new lightpath from an intermediate node to the destination node, e.g., this may not be allowed by administration policy. Moreover, requesting the setup of such a lightpath requires complicated coordination and results in more overhead. To address such constraints, the wavelink needs to differentiate a call as to whether it can request new lightpath setup or has to be carried over an existing lightpath. Thus a more sophisticated wavelink blocking model is needed to handle this situation. A call arriving at a g-link is referred to as an *active call* (with regard to this g-link) if it can request a new lightpath establishment on this g-link to accommodate it, and referred to as a *passive call* if it has to be carried on an existing lightpath on this g-link.

As illustrated in the hierarchy of dynamic traffic grooming in Fig. 13.4, a call arriving at a g-link is routed on a waveroute supporting this g-link. If a call is an active (or passive) call with regard to a g-link, it is correspondingly

referred to as an active (or passive) call on the waveroute supporting this g-link. A waveroute may carry both active and passive calls, since a g-link can carry both types of calls. Actually, the first call carried on a waveroute must be an active call that requests to establish a new lightpath on this waveroute, and later the grooming algorithm can route both active and passive calls from different node-pairs onto this waveroute, which has already become a lightpath. Besides the notations in Section 13.5, let ρ_{nk} and $\tilde{\rho}_{nk}$ denote the offered load of class- $\langle n, k \rangle$ active/passive calls arriving at the wavelink, i.e., service- k active/passive calls coming to the wavelink from waveroute \mathcal{R}_n . The following is a wavelink blocking model developed in [28] to distinguish active and passive calls.

Theorem 2. Let $\hat{\rho}_{nk}(j) = \begin{cases} \rho_{nk} & \text{if } j \leq 0 \\ \rho_{nk} + \tilde{\rho}_{nk} & \text{if } j > 0 \end{cases}$. Then $q(n, j)$ satisfies the recursive relationship

$$j \cdot q(n, j) = \sum_{k=1}^K b_k \cdot \hat{\rho}_{nk}(j - b_k) \cdot q(n, j - b_k), \quad (13.18)$$

where $q(n, j) = 0$ for $j < 0$, $q(n, 0) = q(0)$ and

$$q(0) + \sum_{n=1}^N \sum_{j=1}^C q(n, j) = 1. \quad (13.19)$$

Denote the wavelink blocking probability for class- $\langle n, k \rangle$ active/passive calls as B_{nk} and \tilde{B}_{nk} , respectively. Then

$$B_{nk} = 1 - q(0) - \sum_{j=1}^{C-b_k} q(n, j), \quad (13.20)$$

$$\tilde{B}_{nk} = 1 - \sum_{j=1}^{C-b_k} q(n, j). \quad (13.21)$$

Proof. See [28].

For the end-to-end analysis, we can define the following corresponding traffic loads and blocking probabilities for active/passive calls, similar to the notations in Section 13.6.

- C^{tk}, \tilde{C}^{tk} loads of service- k active/passive calls carried on g-link t .
- $\mathcal{B}^{tk}, \tilde{\mathcal{B}}^{tk}$ blocking probabilities of g-link t for service- k active/passive calls.
- A^{tk}, \tilde{A}^{tk} loads of service- k active/passive calls offered to g-link t .
- $L_k^{tu}, \tilde{L}_k^{tu}$ blocking probabilities of waveroute \mathcal{R}_{tu} for service- k active/passive calls.

$E_k^{tu}, \tilde{E}_k^{tu}$ loads of service- k active/passive calls carried on waveroute \mathcal{R}_{tu} (i.e., class- $\langle t, u, k \rangle$ calls).
 $B_w^{tuk}, \tilde{B}_w^{tuk}$ blocking probabilities of class- $\langle t, u, k \rangle$ active/passive calls on wavelink w .
 $D_w^{tuk}, \tilde{D}_w^{tuk}$ loads of class- $\langle t, u, k \rangle$ active/passive calls offered to wavelink w .

The traffic loads of service- k active/passive calls carried on g-link t can be calculated as follows. Let \mathcal{O}_t denote the set of node-pairs whose calls are active calls with regard to g-link t , and $\mathcal{Q}^t = \{(s, j) \mid s \in \mathcal{O}_t, \text{ and } \mathcal{P}_{s,j} \text{ traverses g-link } t\}$ denote the indices of the paths that offer active calls to g-link t . Correspondingly, we denote $\tilde{\mathcal{Q}}^t$ as the indices of the paths that offer passive calls to g-link t . Then the g-link t carried loads of service- k active and passive calls are

$$\begin{aligned} C^{tk} &= \sum_{(s,j) \in \mathcal{Q}^t} \theta_k^{sj}, \\ \tilde{C}^{tk} &= \sum_{(s,j) \in \tilde{\mathcal{Q}}^t} \theta_k^{sj}. \end{aligned}$$

The traffic loads for both active and passive calls on g-links, waveroutes, and wavelinks are calculated similarly as in Section 13.6, based on C^{tk}, \tilde{C}^{tk} . Then using the extended wavelink blocking model, we can obtain the corresponding wavelink blocking probabilities for active/passive calls. The waveroute and g-link blocking probabilities for active and passive calls are computed similarly as in Section 13.6. We define \mathcal{B}_t^{sk} as the blocking probability of g-link t for service- k calls of node-pair s , which can be computed as below.

$$\mathcal{B}_t^{sk} = \begin{cases} \mathcal{B}^{tk} & \text{if calls of node-pair } s \text{ are active calls on g-link } t \\ \tilde{\mathcal{B}}^{tk} & \text{if calls of node-pair } s \text{ are passive calls on g-link } t \end{cases}$$

The path blocking probability G_k^{sj} in (13.11) is correspondingly computed as follows

$$G_k^{sj} = 1 - \prod_{t \in \mathcal{P}_{s,j}} (1 - \mathcal{B}_t^{sk}). \quad (13.22)$$

Other blocking probabilities are calculated the same as before. The fixed-point equation formed by the nonlinear dependence of the blocking probabilities can be solved similarly and the call loss probabilities between node-pairs can then be obtained.

13.7.3 Arbitrary Alternate Routing

Both the client call routing on the g-link topology and the lightpath routing in the optical network may utilize arbitrary alternate routing instead of link-disjoint routing. The challenge in adopting arbitrary alternate routing is the *repeated g-links/wavelinks* that are present in more than one path/waveroute,

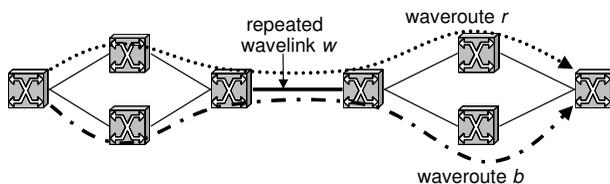


Fig. 13.11. Repeated wavelink

due to the overlapping of several paths between the same node-pair or several waveroutes supporting the same g-link, illustrated in Fig. 13.11. The repeated wavelinks/g-links make individual waveroute/path blocking probabilities not independent from other waveroutes supporting the same g-link or other paths between the same node-pair. For instance, waveroutes r and b in Fig. 13.11 have one repeated wavelink w . If a call offered to waveroute r is blocked due to wavelink w (because there is no available bandwidth), then clearly this call is also blocked on waveroute b . This makes it difficult to compute the joint waveroutes/paths blocking probability. A possible approach to address this issue is to use conditional probability law to consider the conditional joint blocking probability for each given state of the repeated wavelinks/g-links, and then use the total probability law to obtain the unconditional joint blocking probability (see [28]).

13.7.4 Random Wavelength Assignment and Random Traffic Grooming

The work in [29] discussed the analysis of random traffic grooming in optical networks with random wavelength assignment. In random traffic grooming, each client node uses one or more fiber routes to route calls to each destination. An incoming call first tries the single-hop path on a fiber route, and then randomly tries one of the two-hop g-link paths that contains only one intermediate client node on this fiber route. Figure 13.12 illustrates a fiber route from node S to D . There are three two-hop paths on this fiber route in the form of $\{S, X, D\}$, with X being A , B , or C . Two such paths are illustrated in Fig. 13.12. The analysis of this algorithm in [29] takes a similar hierarchical approach

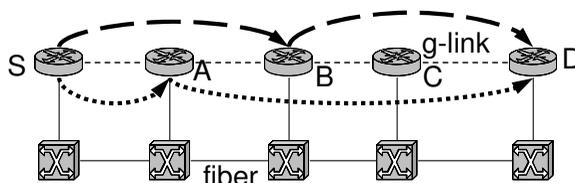


Fig. 13.12. Random traffic grooming

as in Section 13.6, with the hierarchy “single wavelength link (SWL)”–“SWL path”–lightpath–route–path, which is essentially the same hierarchy as wavelink–waveroute–g-link–path–“paths-group on a fiber route” using the terminologies of this chapter. With the random wavelength assignment, the traffic loads on the waveroutes are computed from the g-link offered load differently. The analysis model in [29] partitions the g-link offered load equally on to each waveroute. This model also divides the wavelink into *channels*, similar to time slots, and requires a client call take the same channel(s) on the entire waveroute, which is analyzed through a similar technique as analyzing the lightpath blocking in optical network without wavelength conversion (e.g., see [11]).

13.8 Conclusion

In this chapter, we first discussed the challenges for analyzing dynamic traffic grooming. We addressed these challenges by transforming traffic grooming into a problem of client call routing on a g-link topology. Then we presented an analysis framework that uses a hierarchy to help derive the client call blocking probability from the wavelink blocking probability. Next, a performance model for dynamic traffic grooming was introduced together with a wavelink blocking model that plays a key role in the performance model. At last, we demonstrated how to extend this basic model to various models for accommodating different features and requirements, e.g., wavelength conversion, traffic grooming policy, arbitrary alternate routing, and random traffic grooming.

A major research trend on performance models for dynamic traffic grooming is to enhance them to be used for studying optical network design problems with consideration of traffic grooming, e.g., topology design, routing optimization, and converter placement. Another possible future direction is to consider mixed traffic, e.g., mixed static, scheduled, and dynamic traffic, or mixed sub-wavelength, wavelength, and coarser granularity traffic, and address optical networks with multi-granularity switching, e.g., with both wavelength and waveband switching.

References

1. E. C. Rosen, A. Viswanathan, and R. Callon. (2001, Jan.) Multi-protocol label switching architecture. IETF RFC 3031. [Online]. Available: <http://www.ietf.org>
2. K. Zhu and B. Mukherjee, “Traffic grooming in an optical WDM mesh network,” *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 1, pp. 122–133, Jan. 2002.
3. J. Q. Hu and B. Leida, “Traffic grooming, routing, and wavelength assignment in optical WDM mesh networks,” in *Proc. IEEE Infocom*, pp. 495–501, 2004.

4. C. Xin, Y. Ye, S. Dixit, and C. Qiao, "An integrated lightpath provisioning approach in mesh optical networks," in *OFC Technical Digest*, pp. 547–549, 2002.
5. K. Zhu and B. Mukherjee, "Online approaches for provisioning connections of different bandwidth granularities in WDM mesh networks," in *OFC Technical Digest*, pp. 549–551, 2002.
6. H. Zhu, H. Zang, K. Zhu, and B. Mukherjee, "A novel generic graph model for traffic grooming in heterogeneous WDM mesh networks," *IEEE/ACM Transactions on Networking*, vol. 11, no. 2, pp. 285–299, Apr. 2003.
7. K. Zhu, H. Zhu, and B. Mukherjee, "Traffic engineering in multigranularity heterogeneous optical WDM mesh networks through dynamic traffic engineering," *IEEE Network*, vol. 17, no. 2, pp. 8–15, Mar./Apr. 2003.
8. W. Yao and B. Ramamurthy, "A link bundled auxiliary graph model for constrained dynamic traffic grooming in WDM mesh networks," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 8, pp. 1542–1555, Aug. 2005.
9. C. Ou, K. Zhu, H. Zang, L. Sahasrabudde, and B. Mukherjee, "Traffic grooming for survivable WDM networks—shared protection," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 9, pp. 1367–1383, Nov. 2003.
10. K. Zhu, H. Zang, and B. Mukherjee, "A comprehensive study on next-generation optical grooming switches," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 7, pp. 1173–1186, Sept. 2003.
11. A. Birman, "Computing approximate blocking probability for a class of all-optical networks," *IEEE Journal on Selected Areas in Communications*, vol. 14, no. 5, pp. 852–857, June 1996.
12. A. Sridharan and K. Sivarajan, "Blocking in all-optical networks," in *Proc. IEEE Infocom*, pp. 990–999, 2000.
13. A. Mokhtar and M. Azizoglu, "Adaptive wavelength routing in all-optical networks," *IEEE/ACM Transactions on Networking*, vol. 6, no. 2, pp. 197–206, Apr. 1998.
14. S. Subramaniam, M. Azizoglu, and A. Somani, "All-optical networks with sparse wavelength conversion," *IEEE/ACM Transactions on Networking*, vol. 4, no. 4, pp. 544–557, Aug. 1996.
15. R. Barry, "Model of blocking probability in all-optical networks with and without wavelength changer," *IEEE Journal on Selected Areas in Communications*, vol. 14, no. 5, pp. 858–867, June 1996.
16. Y. Zhu, G. Rouskas, and H. Perros, "A path decomposition approach for computing blocking probabilities in wavelength-routing networks," *IEEE/ACM Transactions on Networking*, vol. 8, no. 6, pp. 747–762, Dec. 2000.
17. R. Srinivasan and A. K. Somani, "A generalized framework for analyzing time-space switched optical networks," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 1, pp. 202–215, Jan. 2002.
18. R. Srinivasan and A. K. Somani, "Analysis of optical networks with heterogeneous grooming architectures," *IEEE/ACM Transactions on Networking*, vol. 12, no. 5, pp. 931–943, Oct. 2004.
19. V. R. Konda and T. Y. Chow, "Algorithm for traffic grooming in optical networks to minimize the number of transceivers," in *Proc. IEEE Workshop on High Performance Switching and Routing*, pp. 218–221, 2001.
20. D. Li, Z. Sun, X. Jia, and S. Makki, "Traffic grooming for minimizing wavelength usage in WDM networks," in *Proc. IEEE ICCN*, pp. 460–465, 2002.

21. C. Xin, C. Qiao, and S. Dixit, "Traffic grooming in mesh WDM optical networks — performance analysis," *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 9, pp. 1658–1669, Nov. 2004.
22. F. P. Kelly, "Blocking probabilities in large circuit switched networks," *Advances in Applied Probability*, vol. 18, pp. 473–505, 1986.
23. S. P. Chung, A. Kashper, and K. W. Ross, "Computing approximate blocking probabilities for large loss networks with state-dependent routing," *IEEE/ACM Transactions on Networking*, vol. 1, no. 1, pp. 105–115, Feb. 1993.
24. A. Girard, *Routing and Dimensioning in Circuit-Switched Networks*. Addison-Wesley, Nov. 1990.
25. J. Bannister, J. Touch, A. Willner, and S. Suryaputra, "How many wavelengths do we really need? a study of the performance limits of packet over wavelengths," *Optical Networks Magazine*, vol. 2, pp. 1–12, Apr. 2000.
26. S. P. Chung and K. W. Ross, "Reduced load approximations for multi-rate loss networks," *IEEE Transactions on Communications*, vol. 41, no. 8, pp. 1474–1481, Aug. 1993.
27. C. Xin, J. Li, X. Cao, and B. Wang, "A route segment technique for blocking analysis of dynamic traffic grooming," in *Proc. IEEE BroadNets*, the 2nd IEEE/Create-Net International Workshop on Traffic Grooming, 2005.
28. C. Xin, "Blocking analysis of dynamic traffic grooming in mesh WDM optical networks," *IEEE/ACM Transactions on Networking*, vol. 15, no. 3, pp. 721–733, June 2007.
29. W. Yao, M. Li, and B. Ramamurthy, "Performance analysis of sparse traffic grooming in WDM mesh networks," in *Proc. IEEE ICC*, pp. 1766–1770, 2005.

Multipoint Traffic Grooming

Ahmed E. Kamal, Ahmad N. Khalil

14.1 Introduction

Several of the new and emerging applications using high-performance networks use one of the multipoint service modes. Under multipoint service, and in the same session, there may be multiple sources, multiple destinations, or both. Multipoint communication can take one of the following forms [1]:

1. *One-to-Many or Multicast*: This type of service is very well known, and several applications belong to this class. These include document distribution, on-demand video distribution, network news distribution, and file distribution and caching.
2. *Many-to-One*: This type of service corresponds to data delivered from a group of users to a single destination. Applications of this type include resource discovery, data collection at a central location, auctions, group polling, and accounting.
3. *Many-to-Many*: In this type of service, several users interact together. Applications requiring this type of traffic include the combination of one-to-many and many-to-one interactions cited above between a speaker and a group of users. They also include multimedia conferencing, synchronized resources, distance learning, distributed simulations, and collaborative processing.

In this chapter, we address the problem of grooming multipoint sessions. Since this area of research is very young, very few results are available. We therefore only consider the grooming of traffic generated in multicast and many-to-one sessions.

This research was supported in part by grant CNS-0626741 from the National Science Foundation.

14.2 Multicast Traffic Grooming

Providing multicast service on optical networks was traditionally considered in two domains, namely, optical multicasting on wavelength routing networks, where a session uses a full wavelength capacity [2, 3], and multicasting on passive star optical couplers in broadcast and select networks [4]. In this section, we consider multicasting on second generation optical networks, i.e., wavelength routed networks, where signals from different multicast sessions have to be groomed. To support this type of service, traffic may have to be duplicated. Electronic equipment functionalities are therefore different, and network design, as well as session provisioning must take this into account. Techniques developed for multicasting in optical networks, which use pure optical multicasting are not applicable to multicast traffic grooming. This is due to at least two reasons: (1) since there is no traffic grooming involved in optical multicasting, and (2) since traffic duplication is implemented in the optical domain using optical splitters, which applies to all traffic on the wavelength channel, even if no branching is required by several sessions. We illustrate the concept of traffic grooming, and the importance of optimally provisioning multicast sessions, using the examples in Fig. 14.1 in which three multicast sessions have to be provisioned over a bidirectional ring with six nodes, A through F. The multicast sessions are given by:

- Session 1: Source = A; Destination = {B, C}; Traffic demand =1 unit;
- Session 2: Source = B; Destination = {C}; Traffic demand = 2 unit;
- Session 3: Source = A; Destination = {F}; Traffic demand = 1 unit;

Each wavelength on the ring can accommodate two traffic units. The two Figures illustrate two different ways of accommodating the sessions, which result in different levels of resource requirements. For example, in Fig. 14.1.(a), 7 units of Line Terminating Equipment (LTE) and two wavelengths are used, while the routing in Fig. 14.1.(b) costs only 6 LTEs and one wavelength.

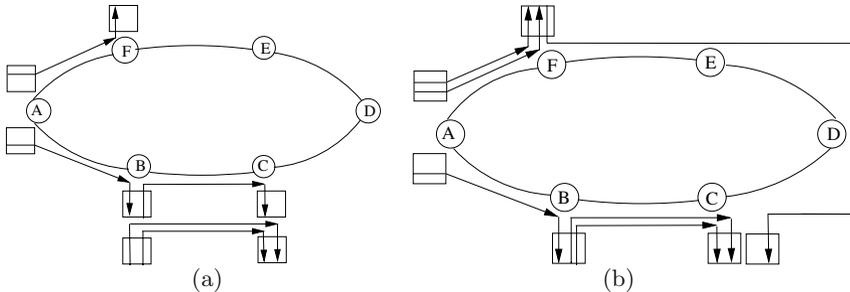


Fig. 14.1. Two different ways of multicast traffic grooming on a ring network

14.2.1 Enabling Technology

When multicast traffic grooming is involved, it may happen that at a node in the network, some of the tributaries aggregated on a certain wavelength need to be duplicated, while others need not be duplicated. In this case, it is natural to use an approach in which the optical signal is terminated at an LTE, and the tributaries are accessed. Tributaries that need to be copied are then duplicated in the electronic domain. The LTE shown in Fig. 14.2 performs this operation.

The implementation of the LTEs in Fig. 14.2 is more suitable when the LTEs are SONET ADMs. However, when using switching nodes supporting Data over SONET (DoS) in Next Generation SONET (NGS) networks [6], e.g., the DoS node shown in Fig. 14.3, two issues will be different. First, an integrated STM/packet switching architecture is used, hence obviating the need for digital cross connects. Second, data duplication may not be needed since data is encapsulated into Generic Framing Procedure (GFP) frames, and frames may be transmitted on different ports, provided that the overhead is adjusted to reflect the new destination. Reference [7] introduced a multicast extension header for GFP.

In references [5, 8], it was argued that implementing some of the traffic duplication in the optical domain might be less expensive, since the cost of (passive) optical splitters is considerably less than the cost of electronic LTEs. This is especially true if the duplicated traffic continues to use the same wavelength on different OXC output ports, and if the LTEs are only used for traffic duplication, and not for traffic dropping at a destination. However, electronic devices are still needed if traffic needs to be added to wavelength channels. Moreover, optical splitting can result in overusage of wavelength channels. Therefore, the authors introduced the node architecture shown in

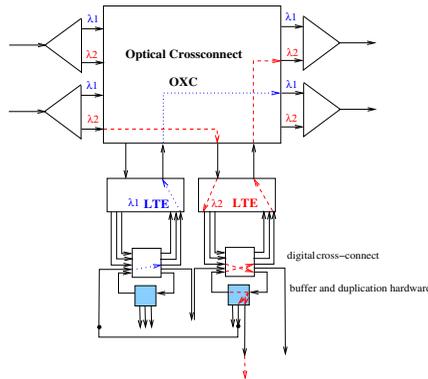


Fig. 14.2. Multicast traffic duplication in the electronic domain

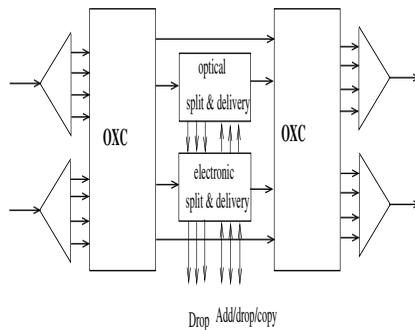


Fig. 14.3. Node architecture employing optical and electronic duplication proposed in [5]

Fig. 14.3 which implements both electronic and optical duplication, based on need and cost. Such nodes are known as translucent nodes.

14.2.2 Network Model

We consider a network with N nodes, where the network graph can assume any arbitrary mesh topology. Each edge in the graph corresponds to a pair of fibers between a pair of nodes, which are used for transmission in two different directions. Each fiber has W wavelength channels. The capacity of each channel is expressed in g , the grooming factor, which is the number of basic units of traffic which can be carried on a channel. For example, if the transmission rate on a wavelength channel is OC-48, or 2.4 Gbps, and the basic transmission unit is OC-3, or 150 Mbps, then $g = 16$. There are K multicast sessions in the network. Each session, a , where $1 \leq a \leq K$, is identified by the tuple $\langle s_a, D_a, m_a \rangle$. s_a is the source of the session, D_a is the set of destinations of the session, and m_a is the number of traffic units to be delivered from s_a to all destinations $d \in D_a$.

14.2.3 Static Multicast Traffic Grooming: Exact Approaches

Integer Linear Program (ILP)

We first consider that all traffic duplication is carried out in the electronic domain, using the node architecture shown in Fig. 14.2. An exact approach for the static multicast traffic grooming problem was presented in [7] assuming this node architecture. We present a simplified version of this approach, in the form of an Integer Linear Program (ILP), which minimizes the number of electronic equipment, LTEs. Table 14.1 presents the basic parameters needed for ILP formulation. The ILP can be formulated as follows:

Table 14.1. Multicast traffic grooming ILP basic parameters

Q	A very large integer number, (in our case $Q \geq N^2 - N$)
F_{mn}	a binary variable that indicates the presence of an edge between nodes m and n , $1 \leq m, n \leq N$: equals 1 if and only if nodes m and n are adjacent
$F_{mn}^{ij,w}$	number of lightpaths between node pair (i, j) routed on fiber (m, n) on wavelength w
T_n	number of LTEs at node n
L_{ij}^w	number of lightpaths from node i to node j on wavelength w
L_{ij}	number of lightpaths from node i to node $j = \sum_w L_{ij}^w$
$Z_{ij}^{a,d}$	a binary indicator: equals 1 if and only if session a , destined to d , is employing a lightpath from i to j as an intermediate virtual link
M_{ij}^a	a binary indicator: equals 1 if and only if $\exists d \in D_a$, such that $Z_{ij}^{a,d} = 1$. This means that session a is using lightpath (i, j) to reach at least one destination.

Objective function:

$$\text{Minimize : } \sum_{1 \leq n \leq N} T_n \quad (14.1)$$

Subject to:

- *Number of LTEs:*

$$T_i \geq \sum_w \sum_{j, j \neq i} L_{ij}^w \quad \forall i \quad (14.2)$$

$$T_i \geq \sum_w \sum_{j, j \neq i} L_{ji}^w \quad \forall i \quad (14.3)$$

- *Lightpath level constraints:*

$$\sum_{m, \text{for } F_{mi}=1} F_{mi}^{ij,w} = 0 \quad \text{and} \quad \sum_{n, \text{for } F_{in}=1} F_{jn}^{ij,w} = 0 \quad \forall i, j, w \quad (14.4)$$

$$\sum_{m, \text{for } F_{mx}=1} F_{mx}^{ij,w} = \sum_{n, \text{for } F_{xn}=1} F_{xn}^{ij,w} \quad \forall w, i, j, x; x \neq i, j \quad (14.5)$$

$$\sum_{m, \text{for } F_{mj}=1} F_{mj}^{ij,w} = L_{ij}^w \quad \text{and} \quad \sum_{n, \text{for } F_{in}=1} F_{in}^{ij,w} = L_{ij}^w \quad \forall i, j, w \quad (14.6)$$

$$\sum_i \sum_{j, \text{for } F_{mn}=1} F_{mn}^{ij,w} \leq 1 \quad \forall m, n, w \quad (14.7)$$

and

$$\sum_w L_{ij}^w = L_{ij} \quad \forall i, j \quad (14.8)$$

The above constraints are self-explanatory, and they make sure that the traffic flows on each lightpath are conserved. In addition, they evaluate the number of lightpaths between node pair (i, j) , viz., L_{ij} . The constraints apply if the traffic is either multipoint, or point-to-point.

- *Multicast session topology constraints:*

$$\sum_{i, i \neq s} Z_{is}^{a,d} = 0 \quad \text{and} \quad \sum_{j, j \neq d} Z_{dj}^{a,d} = 0 \quad \forall a, d \in D_a \quad (14.9)$$

$$\sum_{j, j \neq s} Z_{sj}^{a,d} = 1 \quad \text{and} \quad \sum_{i, i \neq d} Z_{id}^{a,d} = 1 \quad \forall a, d \in D_a \quad (14.10)$$

$$\sum_{i, i \neq x} Z_{ix}^{a,d} = \sum_{j, j \neq x} Z_{xj}^{a,d} \quad \forall a, d \in D_a, x, (x \neq s, d) \quad (14.11)$$

The above constraints are similar to the lightpath constraints, except that they apply to the multicast tree traffic, and the continuity is over the lightpaths, and not on the physical links.

Since two or more multicast sessions in the same destination set can share a lightpath to reach their respective destinations, the bandwidth along the shared path could be shared too. The above constraints ensure that the bandwidth of all lightpaths from i to j have not exceeded the physical capacity. The use of the M_{ij}^a variables will avoid multiple counting of the same bandwidth m_a used by several destinations of the same session, and on the same lightpath.

$$M_{ij}^a \geq \sum_{d \in D_a} Z_{ij}^{a,d} / Q \quad \forall a, i, j \quad (14.12)$$

$$M_{ij}^a \leq \sum_{d \in D_a} Z_{ij}^{a,d} \quad \forall a, i, j \quad (14.13)$$

- *Capacity constraints*

$$\sum_{a=1}^K m_a M_{ij}^a \leq L_{ij} * g \quad \forall i, j \quad (14.14)$$

The above ILP produces the locations of the LTEs, and also provisions sessions, in terms of routes and wavelength assignment.

Other Exact Approaches

Other exact approaches were introduced in [5, 9] and [8]. The model in [9] considered multicast traffic grooming on a mesh network, where the objective was to reduce the number of wavelength links. However, the model in [9] assumed that multicast sessions were routed in advance, which removes the routing problem from the formulation. In addition, wavelength conversion

was used, which removes the constraints on wavelength continuity. Moreover, optical splitting can be used. The problem therefore reduced to a bin packing problem on one link at a time.

Using the translucent node architecture shown in Fig. 14.3, the authors in [5] introduced an exact approach, in the form of an Integer Nonlinear Programming formulation, for the network design and session provisioning. The objective function was to also minimize the cost of the electronic equipment. The formulation bears resemblance to the above one, except in one aspect, which is used to determine whether or not an add/drop port is required at a node. This involves terms, for each wavelength, and for each pair of links having a common node, which are the products over all sessions of other terms which are 1 if an add/drop is not needed for each of these sessions. Using a similar architecture, the authors in [8] also introduced an ILP formulation for multicast traffic grooming on ring networks with one traffic unit per session.

The above multicast models assume that all destinations within a destination set have the same traffic requirements. However, different modes of operation of multicasting service include:

- Multicasting with *partial receiver reachability*, in which a subset of the receivers must receive the traffic originating from the source, while every attempt should be made to deliver the traffic to the remaining receivers, if possible, and without increasing the network cost. This corresponds to applications in which users are prioritized, and lower priority users can be accommodated only if this imposes no additional cost. It also corresponds to the case in which the network is designed under tight budget constraints, and users will have to be prioritized in terms of which user is to be served under the given constraints.
- Multicasting in which *traffic may be pruned, or thinned* while propagating downstream, depending on the receivers' needs. Applications of this type include multi-layer coded video streams, in which different receivers may require different stream qualities.

The above two cases were also considered in [7], in which the above ILP was modified in order to accommodate these two cases. The destination set for session a , D_a , is partitioned into two classes, D'_a , and D''_a , such that $D_a = D'_a \cup D''_a$, and $D'_a \cap D''_a = \phi$. All destinations $d' \in D'_a$ have a traffic requirement of m'_a , while destinations $d'' \in D''_a$ have a traffic requirement of m''_a .

For the partial destination reachability case, $m'_a = m''_a$, but D''_a can be reached only if this is not going to increase the network cost. The objective function in this case becomes:

$$\text{Minimize} : \alpha \cdot \sum_n T_n - \beta \cdot \sum_a \sum_{d \in D''_a} \sum_{j, j \neq s} Z_{sj}^{a,d} \quad (14.15)$$

The weights in the above objective function, α and β , should be chosen such that $\alpha \gg \beta$.

Also, for the traffic thinning case, $m'_a \neq m''_a$, and usually $m'_a > m''_a$. The objective function remains the same as in Equation (14.1). However, the constraints must be revised by including a binary variable for each session a on each lightpath, (i, j) , in order to select the correct traffic level, depending on the requirements of the downstream destinations. That is, the traffic delivered on a lightpath used by a session, is by default equal to the m''_a . If the lightpath is used to deliver a traffic level of m'_a , which will cause a binary variable to be equal to 1, a constant quantity equal to $m'_a - m''_a$ is added through multiplying it by the above binary variable.

14.2.4 Static Multicast Traffic Grooming: Heuristic Approaches

Unicast traffic grooming is known to be an NP-hard problem [10, 11]. Multicast traffic grooming is therefore an even harder problem. Therefore, the above exact approaches are only useful for use with small networks, and with limited traffic levels. For large networks, and for higher levels of traffic, other approximate approaches are therefore needed. In addition, the exact approach is still useful for assessing the accuracy of the approximate approaches. Several heuristic approaches were proposed in the literature and will be reviewed here.

Grooming with Lightpath Replacement (GLR)

Reference [7] introduced a heuristic approach based on some observations made from the exact solution of small sized examples. It was found that many sessions are routed along non-shortest paths, and several destinations were reached through lightpaths carrying traffic for such sessions. Those lightpaths were provisioned in a way that makes use of LTEs which were already in place. This observation can be made by inspecting the examples in Figure 14.1.

To illustrate the concept of exploring other trees which can reduce the number of LTEs, consider the example in Fig. 14.4.(a). Two lightpaths were established to deliver traffic (which belong to different sessions) from nodes A to B and C. The grooming factor, g , is 2 units, and the bandwidth requirement per destination is 1 unit. This results in using 4 LTEs. However, when remov-

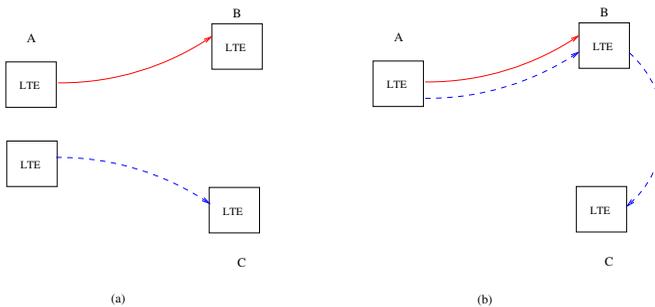


Fig. 14.4. An example to show the two steps of the heuristic

ing the link (A,C), C can be serviced using two branches from A to B, and then from B to C, hence requiring only 3 LTEs, as shown in Figure 14.4.(b).

Since the search space can be very large, the step of exploring other routes and wavelengths is performed by inspecting one lightpath at a time, removing one link on the lightpath, and inspecting whether using existing LTEs can lead to reaching the affected destination while using fewer LTEs.

Heuristics Using Translucent Node Architecture

Reference [5] introduced three heuristics. Although the heuristics have been introduced for networks using a combination of optical and electronic duplication, the algorithms can still work with electronic duplication only.

k-Shortest Path Trees (*k*-SPT)

The first algorithm is the *k*-Shortest Path Tree (*k*-SPT) where the *k*-SPTs are constructed by constructing the SPT, and then removing the links from this SPT, one at a time, in order to obtain the remaining $k - 1$ SPTs.

An example is shown in Fig. 14.5, where there are two sessions: the first one is $\langle A, \{C, G\}, 1 \rangle$, while the second one is $\langle A, \{E, G\}, 1 \rangle$. The grooming factor, g , is 2. Figure 14.5.(a) shows the routing under SPT for each session, which requires 10 LTE ports, which are shown as squares. However, in Figure 14.5.(b), optical splitting is used at node B, which reduces the number of LTEs to 5. Note that in Fig. 14.5.(b), traffic from the second session is delivered to node C, which is not part of the destination set of this session. This is due to the optical splitting implemented at node B.

Grooming with Rerouting Sessions (GRS)

The second algorithm in [5] is called Grooming with Rerouting Sessions (GRS). In this algorithm, if a session cannot be routed on its shortest path tree on already used wavelengths due to the unavailability of resources on some links, such links are treated as bottleneck links. If the number of such bottleneck links does not exceed two links, then the session is rerouted to

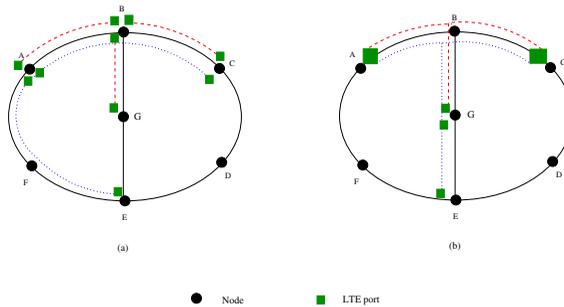


Fig. 14.5. An example of the *k*-SPT heuristic

avoid those links. Otherwise, the next available wavelength is assigned to the shortest path tree.

Grooming by Computing Overlapped Trees (GCOT)

The third and last algorithm is Grooming by Computing Overlapped Trees (GCOT). This algorithm tries to pack sessions onto wavelengths in a greedy manner.

Performance Evaluation

It was shown in [5] that the k-SPT algorithm outperforms both the GRS and GCOT algorithms, both in terms of the number of LTEs and the number of wavelength channels. Also, as the value of k increases, the performance of k-SPT increases further, but of course at the expense of an added complexity. By comparing the performance of the algorithms to the optimal solution obtained by solving the nonlinear program mentioned earlier, it was shown that the results of the algorithm are within 66% of the optimal. However, in most cases they are within 10–20% of the optimal.

The GLR algorithm described above is very much related to the k-SPT algorithm, since it also investigates several shortest paths. However, there are two differences between the two algorithms. First, under GLR, the number of alternate shortest paths is not fixed, while it is fixed under k-SPT. Second, under k-SPT, shortest path trees are constructed after removing a link, while under GLR, the original tree is augmented in order to reach the destinations which have been disconnected.

Multicast Traffic Grooming with Sparse Optical Splitters

Reference [9] considered a network with Multicast Capable (MC) nodes (equipped with optical splitters) and Multicast Incapable (MI) nodes. It is assumed that each MI node has at least one adjacent MC node as well as each node is equipped with a bank of wavelength converters. The paper presented a two phase algorithm for provisioning multicast traffic sessions, namely, *Multicast Tree Construction (MTC)* and *First Fit Grooming Algorithm (FFGA)*. MTC constructs shortest path trees from each MC node, and allows an MI node to use an adjacent MC node as a virtual source, or root. The MI node, therefore, routes its multicast traffic to that adjacent MC node, and the MC node will split the traffic and route it to the destinations. The FFGA is a first fit greedy algorithm that allocates traffic streams on a link to the first available wavelength channel on that link. If no wavelength channel is available, a new wavelength channel is created.

14.2.5 Dynamic Multicast Traffic Grooming

With the use of IP over WDM, MPLS over WDM, or with NGS, traffic sessions now tend to exhibit a dynamic nature, as opposed to the static

nature assumed in the previous section. Sessions arrive according to a certain arrival process, and they are characterized by holding times, which are taken from a certain distribution. Since it is practically impossible to design optical networks such that they accommodate all such dynamic sessions, most efforts have concentrated on devising call acceptance, and session provisioning strategies that will try to reduce the session blocking probabilities.

A number of session provisioning strategies for multicast traffic were introduced in the literature, and will be presented below.

The Maximizing Minimum Freeload (MMFL) Algorithm

Reference [12] introduced a session provisioning strategy for dynamic multicast traffic, assuming translucent nodes with the architecture shown in Fig. 14.3. The objective of this algorithm is to increase resource utilization, and to minimize the blocking probability for future arriving requests. This is done by using paths which will maximize the bandwidth capacity left after routing a multicast tree. In other words, the problem can be considered as a max-min problem, where the minimum left over bandwidth on all links is maximized.

The bandwidth capacity left after routing the tree is called the *freeload*, and the freeload on link (i, j) which has an available capacity $c_{i,j}$ on wavelength w after routing a session with bandwidth requirement B on this link, and this wavelength, is given by

$$\frac{c_{i,j} - B}{g} \quad (14.16)$$

where g is the grooming factor. The minimum freeload on all wavelengths is calculated assuming that the session is provisioned, and the routing that yields the maximum over these minima is used.

The algorithm introduced in [12] is called *Maximizing Minimum Freeload* (MMFL). It is assumed that a freeload graph for each wavelength, w , is formed for each session. The current such graph is referred to as G_w , and the new graph, G'_w , is formed for each session by applying Equation (14.16) on all links of the graph using the session requirement. After deciding on the routing and wavelength assignment of the session, only G_w on which the session is routed is updated by discounting the allocated link capacities. MMFL is executed for every arriving multicast session.

Figure 14.6 shows an example of the MMFL in which there are two sessions, the first is $\langle A, \{C, G\}, 1 \rangle$, while the second is $\langle A, \{C, E\}, 1 \rangle$. Both sessions have bandwidth requirements of one unit, while $g = 2$. It is assumed that there is a single wavelength in the network. In Fig. 14.6.(a), session 1 arrives, and the freeload factor for all fibers is calculated as $\frac{2-1}{2} = 0.5$. Session 1 is

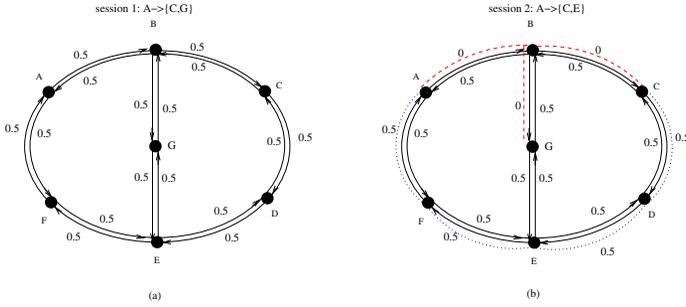


Fig. 14.6. An example of the operation of the MMFL algorithm

routed on the SPT shown by the dashed lines in Figure 14.6.(b). When session 2 arrives, in Fig. 14.6.(b) the freeload factors for session 2 are calculated as shown in the figure. Fibers used by session 1 have a freeload factor of $1 - 1/2 = 0$, while all other fibers have a freeload factor of 1. Session 2 is provisioned on the SPT shown by the dotted lines in Fig. 14.6.(b). If there was another wavelength in the network, the freeload diagram for the second wavelength, when session 2 arrives, would be identical to that in Fig. 14.6.(a), and session 2 would be provisioned similar to that in Fig. 14.5.(b), but on the second wavelength.

It was shown in [12] that MMFL outperforms both fixed and adaptive SPT provisioning techniques in terms of call acceptance ratio, since MMFL attempts to uniformly distribute the session loads over wavelengths, hence resulting in a better chance of call acceptance. The same observation also holds for resource utilization.

It is to be noted that this algorithm does not consider the effect of using optical splitting on accommodating sessions. For example, the route provisioning shown in Fig. 14.5.(b) will not be selected by this algorithm.

Sequential and Hybrid Multicast Traffic Grooming

References [13, 14] assumed a node architecture similar to that in Fig. 14.3, except that duplication is always done in the optical domain and that multicasting is always supported on a light-tree. All traffic on the same light-tree will be delivered to the same set of destinations on the light-tree, whether they are part of the destination set or not.

Multicast Grooming Approaches

In order to support dynamic traffic, once a multicast session is to be accommodated, one of four traffic grooming approaches may be used. These approaches are described below, with the aid of the example given in Fig. 14.7, assuming that there is one wavelength channel and all sessions have a traffic requirement of 1 unit, while $g = 2$. It is also assumed that sessions 1, 2 and 3 are already established on the network.

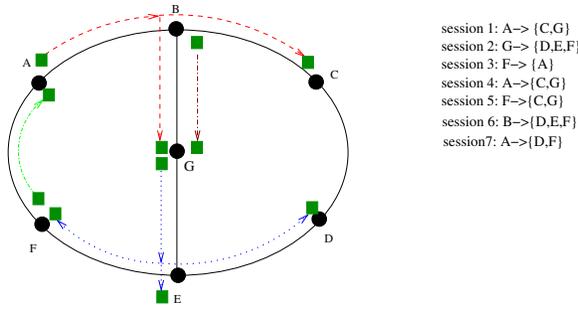


Fig. 14.7. Examples of the approaches used in sequential multicast traffic grooming

Single-Hop Provisioning: In this approach, an existing light-tree with available bandwidth to support the new session is used to groom the multicast session at the logical layer. The new multicast session will be served using a single-hop logical route. Note that for this approach, the new multicast session must have the same source and multicast destinations as those associated with the logical route. For example, in Fig. 14.7, when session 4 arrives, it can be accommodated on the light-tree used by session 1.

Multi-Hop Provisioning: A multicast session can be provisioned on the logical topology by routing data on more than one light-tree (only two hops are allowed here).

In this approach, an existing light-tree whose destinations are the same as those of the new multicast session (called the *to-destinations light-tree* (TDLT)), but with traffic from a different source, is used with combination of a single-hop lightpath whose source is the same as the source of the new session and its destination is the source of the TDLT (called *from-source light-path* (FSLP)). The new multicast session is served on the combination of FSLP and TDLT. The TDLT and FSLP must have enough capacity to accommodate the arriving session. For example, when session 5 arrives, it can be provisioned using an FSLP F to A (which is used to accommodate session 3), and then TDLT from A to C and G (which is used to accommodate session 1).

Hybrid Provisioning: A session can be routed/provisioned over a combination of existing lightpaths or light-trees (logical routing) and a newly created light-path or light-tree (physical routing, e.g., RWA/MC-RWA). In this approach, the optical layer must keep updated databases about the connectivity of both the logical and physical topologies as well as the resource utilization across both layers. The idea of a hybrid provisioning approach is defined to find a combined route of an existing logical segment (light-tree) and an un-provisioned physical segment (lightpath) that needs to be set up.

For example, when session 6 arrives, the light-tree used by session 2 from G to {D,E,F} can be used to reach F, E and D from G. In addition, a light-path from B to G is needed, and is therefore provisioned.

Non-Restricted Multi-Hop Provisioning: This approach allows light-trees to carry traffic destined to only a subset of the destinations of the light-tree (this includes an empty subset, which corresponds to the case in which the light-tree is used as a first hop). For example, when session 7 arrives, the signal can be transmitted on the light-tree used by session 1. Since none of the destinations of session 7 are on this tree, at node G the signal is carried on the light-tree used by session 2 to reach D and F. Notice here that the signals also reach nodes C, and E, which are not in the destination set of session 7.

Traffic Grooming Heuristics

Using the above four approaches, the authors introduced four heuristics for provisioning multicast calls with subwavelength traffic rates.

Logical-First Sequential Routing (LFSR): The network first tries to accommodate the call on the logical topology making use of the already existing connections. Depending on the grooming approach to be used (i.e., Single-Hop or Multi-Hop), if no available single/multi-hop route were found on the logical topology, then a new light-tree destined to all the multicast destinations is set up on the physical topology.

Physical-First Sequential Routing (PFSR): The network attempts to accommodate a session on the physical layer first. If the new light-tree is established successfully, a new logical/virtual route (light-tree) is created in the logical layer. If the physical routing fails, then single or multiphop routing on the logical topology is attempted.

Logical-First Hybrid Routing (LFHR): The hybrid provisioning approach is combined with sequential provisioning approaches to achieve a grooming scheme that is biased toward serving multicast traffic demands in comparison with all other sequential grooming approaches. Specifically, the *LFSR* is invoked to find an existing lightpath to provision an FSLP to deliver the traffic from the source of the session to the root of a TDLT that reaches the destinations. If this fails, a new light-path is created and used to reach the source of the TDLT.

Non-Restricted Logical-first Sequential Routing (NRLFSR): Similar to *LFSR*, but a non-restricted multi-hop provisioning approach is used when *LFSR* fails.

Performance Evaluation

Using a simulation study, and assuming that only one third of the calls are multicast, it was shown in [14] that among the restricted algorithms (*LFSR*, *PFSR* and *LFHR*), *PFSR* has the smallest blocking probability. The use of multihop session provisioning reduces the blocking probability even further,

and provides a reduction in the blocking probability that can be very close to 20%. On the other hand, it was also shown that the non-restricted routing always outperforms constrained routing at light load. The last observation is expected since the routing problem has fewer constraints. However, as the load is increased, the restricted LFSR outperforms NRLFSR, since it is more conservative in using bandwidth to accommodate new sessions. When most of the sessions are multicast sessions, the non-restricted routing approach prevails.

14.2.6 Multicast Traffic Grooming: Summary

In Table 14.2, we provide a summary of the above multicast traffic grooming protocols. For each protocol, the duplication domain for which the algorithm was designed, the mode of operation, and the complexity of the algorithm are indicated.

14.3 Many-to-One Traffic Grooming

In this section, we illustrate the concept of many-to-one traffic grooming, and how aggregating traffic, if possible, can result in resource reduction, with the help of a simple example. Figure 14.8 shows a network consisting of 7 nodes. Suppose that there is only one session which consists of 4 sources, S_1 , S_2 , S_3 and S_4 , located at nodes 1, 3, 4 and 6, respectively. The destination of this session, d , is located at node 7. Let the capacity of a single wavelength be 48 units and the traffic originating from each source be 24 units. We refer to the traffic originating from each source as a stream. Without traffic grooming, the session $\{S_1, S_2, S_3, S_4 \rightarrow d\}$ must be accommodated using 3 wavelengths and 8 LTEs (2 LTE for each source-destination pair). However, with traffic grooming, the number of LTEs and wavelengths can be reduced

Table 14.2. Summary of multicast traffic grooming heuristic algorithms

Algorithm	Duplication domain	Operation	Complexity
GLR [7]	electronic	Static	$O(K \cdot M \cdot N^2 \log_2 N)$
k-SPT [5]	electronic-optical	Static	$O(K \cdot k \cdot N^2 \log_2 N)$
GRS [5]	electronic-optical	Static	$O(K^2 \cdot M \cdot N^2 \log_2 N)$
GCOT [5]	electronic-optical	Static	$O(K^2 \cdot M \cdot W)$
MTC+FFGA [9]	optical	Static	$O(K \cdot N^2 \log_2 N + K \cdot M \cdot W)$
MMFL [12]	electronic-optical	Dynamic	$O(K \cdot N^2 \log_2 N)$
LFSR [14]	optical	Dynamic	$O(W \cdot N^2)$
PFSR [14]	optical	Dynamic	$O(W \cdot N^2)$
LFHR [14]	optical	Dynamic	$O(W \cdot N^2)$
NRLFSR [14]	optical	Dynamic	$O(W \cdot N^2)$

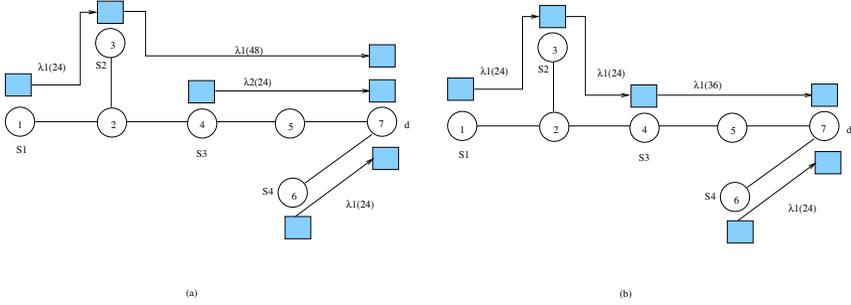


Fig. 14.8. An example of many-to-one traffic grooming: (a) without aggregation and traffic reduction; (b) with aggregation and traffic reduction

to 7 and 2, respectively, as shown in Fig. 14.8(a). If the streams from different sources destined to the same destination, d , may be aggregated on their way, the resources can be further reduced. If we assume that every time traffic is aggregated, 50% of the aggregated traffic will be dropped, then streams from S_1 and S_2 are first aggregated at node 3 and then streams from S_1 , S_2 and S_3 are aggregated at node 4. This accommodation requires only 1 wavelength and 6 LTEs, as shown in Fig. 14.8.(b). Therefore, many-to-one traffic grooming with aggregation can result in a substantial reduction in the cost of the network in terms of the number of LTEs and the number of wavelengths.

14.3.1 Network Model

Similar to the multicast traffic grooming case, we consider a network with N nodes, and an arbitrary mesh topology. Each link corresponds to two bidirectional fibers in opposite directions, where each fiber carries W wavelengths, and the grooming ratio is g . We allow each many-to-one session to traverse multiple lightpath hops, while each lightpath hop itself may span multiple physical links.

Each session, a , will have one destination, d_a , and a set of sources, S_a , with each source $s \in S_a$ generating a number of basic traffic units, $m_{a,s}$. The stream from s to d in session a is denoted by $\phi_{s,d,a}$. Traffic from different sources in the same session may be aggregated such that if f streams from session a are aggregated, each stream contributes a fraction r_f^a of its traffic.

14.3.2 Static Many-to-One Traffic Grooming with Aggregation

We address the many-to-one traffic grooming problem such that, given a traffic matrix, all the traffic should be accommodated with the least number of LTEs. We also consider the situation in which traffic can be aggregated using arbitrary, but application dependent, aggregation ratios. Aggregation

can be used to reduce the amount of traffic that has to be carried by the network, and hence possibly reduce the amount of resources. We assume that aggregation, when performed, will be implemented at layer 1. We expect that the most flexible way of implementing many-to-one traffic grooming would be by using DoS switches, and NGS protocols, e.g., by delineating and flagging data blocks which may be dropped during aggregation. GFP extension headers [15] can facilitate this implementation. This area of research is very new, and all the approaches presented in this section are based on reference [16].

Exact Approach: Mixed Integer Linear Program (MILP) Formulation

The optimal formulation for the network and connection provisioning problem, such that the number of LTEs is minimized, is based on a Mixed Integer Linear Program (MILP). Several of the variables and constraints defined for the multicast traffic grooming problem of Section 14.2.3 are valid for the many-to-one traffic grooming problem. In addition to the parameters defined in the network model above, and those defined in Table 14.3, the MILP uses the following parameters:

Objective function:

$$\text{Minimize : } \sum_n T_n \quad (14.17)$$

Subject to:

- *Number of LTEs:*
The constraints are similar to those in Equations (14.2) and (14.3).
- *Lightpath level constraints:*
The constraints are also similar to those in Equations (14.4), (14.5), (14.6), (14.7) and (14.8).

Table 14.3. Many-to-one traffic grooming MILP basic parameters

$Z_{ij}^{a,s}$	a real number between 0 and 1, which takes non zero values if and only if the traffic stream from source $s \in S_a$, is using a lightpath from i to j
M_{ij}^a	a binary indicator; is 1 if and only if at least one of the sources of session a is using a lightpath from node i to j to reach the destination, i.e., $\exists s \in S_a$, such that $Z_{ij}^{a,s} = 1$
$I_{ij}^{c,a,f}$	a binary indicator; is 1 if and only if the number of sources $s \in S_{c_a}$, on a lightpath from node i to j , is $\geq f$
X_{ij}^a	a real number which represents the amount of traffic on a lightpath from node i to j due to all sources in S_a
$J_{ij}^{a,b}$	a binary indicator; is 1 if and only if session a and b are groomed on the same lightpath from i to j
$Y_{ij}^{a,b}$	a real number and is a product of $J_{ij}^{a,b}$ and X_{ij}^b

- *Session topology constraints:*

$$\sum_i Z_{is}^{a,s} = \sum_j Z_{dj}^{a,s} = 0 \quad \forall a, s \in S_a \quad (14.18)$$

$$\sum_{j,j \neq s} Z_{sj}^{a,s} = \sum_{i,i \neq d} Z_{id}^{a,s} = 1 \quad \forall a, s \in S_a \quad (14.19)$$

$$\sum_{i,i \neq x, i \neq d} Z_{ix}^{a,s} = \sum_{j,j \neq x, j \neq s} Z_{xj}^{a,s} \quad \forall a, s \in S_a, x, (x \neq s, d) \quad (14.20)$$

The above constraints conserve traffic flows on the lightpaths from the sources to the destination, in each of the sessions.

Moreover, the following constraints will ensure that once data streams, from more than one source, are aggregated at some node, they will not split again until they reach the destination.

$$M_{ij}^a \geq \sum_{s \in S_a} Z_{ij}^{a,s} / Q \quad \forall a, i, j \quad (14.21)$$

$$M_{ij}^a \leq \sum_{s \in S_a} Z_{ij}^{a,s} \quad \forall a, i, j \quad (14.22)$$

$$\sum_j M_{ij}^a \leq 1 \quad \forall a, i \quad (14.23)$$

To model aggregation, and to determine the amount of traffic before and after aggregation on each of the lightpaths, the following constraints are used to set the binary indicator $I_{ij}^{a,f}$ on the lightpaths from i to j to one if and only if the number of sources $s \in S_a$ is greater than or equal to f .

$$\sum_{s \in S_a} Z_{ij}^{a,s} = \sum_{f=1}^{|S_a|} I_{ij}^{a,f} \quad \forall a, i, j \quad (14.24)$$

$$I_{ij}^{a,f} \leq I_{ij}^{a,f-1} \quad 2 \leq f \leq |S_a|, \forall a, i, j \quad (14.25)$$

Once the $I_{ij}^{a,f}$ variables have been determined by the above constraints, they are used to determine the exact amount of traffic after aggregation on lightpath(s) from node i to node j due to the sources in S_a . This is done using the following set of constraints.

$$X_{ij}^a \geq r_f^a * \sum_{s \in S_a} m_{a,s} Z_{ij}^{a,s} - Q \sum_{s \in S_a} Z_{ij}^{a,s} + Q * f * I_{ij}^{a,f} \quad 1 \leq f \leq |S_a|, \forall a, i, j \quad (14.26)$$

$$X_{ij}^a \leq r_f^a * \sum_{s \in S_a} m_{a,s} Z_{ij}^{a,s} + Q \sum_{s \in S_a} Z_{ij}^{a,s} - Q * f * I_{ij}^{a,f} \quad 1 \leq f \leq |S_a|, \forall a, i, j \quad (14.27)$$

The above constraints will compute the exact amount of traffic after aggregation on lightpath(s) from node i to node j due to the sources in S_a , since when $f < \sum_{s \in S_a} Z_{i,j}^{a,s}$ or $f > \sum_{s \in S_a} Z_{i,j}^{a,s}$, Equations (14.26) and (14.27) will evaluate X_{ij}^a to be in the range $[-Q, +Q]$. However, when $f = \sum_{s \in S_a} Z_{i,j}^{a,s}$, X_{ij}^a will evaluate to the correct amount of traffic after aggregation.

- *Capacity constraint:*

The following constraint bounds the total capacity on lightpath(s):

$$\sum_{a=1}^K X_{ij}^a \leq L_{ij} * g \quad \forall i, j \quad (14.28)$$

Similar to the multicast traffic grooming problem, the above MILP will determine the locations of the LTEs, as well as the provisioning of sessions.

Heuristic Approach

Reference [16] has also presented a heuristic approach for many-to-one traffic grooming on random topologies, which also performs data aggregation. The heuristic is based on a Dynamic Programming (DP) style algorithm, which is referred to as the Many-to-one Traffic Grooming heuristic based on Dynamic Programming (MTG-DP). That is, the algorithm proceeds in stages, where in stage i different subsets of i streams each are routed and provisioned. The number of subsets, as will be explained below, is equal to the total number of streams. In stage i , stream j is provisioned together with $i - 1$ streams, where such streams are chosen as one of the subsets of streams provisioned in stage $i - 1$, such that the provisioning cost is minimized. This makes the computations at stage i dependent on stage $i - 1$ only, hence the dynamic programming style.

An example of the DP-style algorithm is shown in Fig. 14.9. If each many-to-one session, a , consists of $|S_a|$ streams, one from each source $s \in S_a$, and if Φ is the total number of streams from all the many-to-one sessions, i.e., $\Phi = \sum_a |S_a|$, then, as shown in the figure, the algorithm consists of Φ stages, while each stage further consists of Φ steps. At each step of a stage, a new stream is accommodated (i.e., the stream is assigned a specific route and a specific wavelength) after conducting a systematic search, such that the decisions at each step of a stage depend only on the previous stage. For the systematic search at each stage, all the assignments of the previous stage are examined and the assignment that can accommodate the current stream with the least cost is determined. In Fig. 14.9, steps 2 and 4 of stages 2 and 3 are shown in detail. The outcome of each step is a selection of a set of accommodated streams, which are called an assignment. At the end of each stage, a collection of several best assignments are available, and after the final stage, the assignment with the least cost is selected.

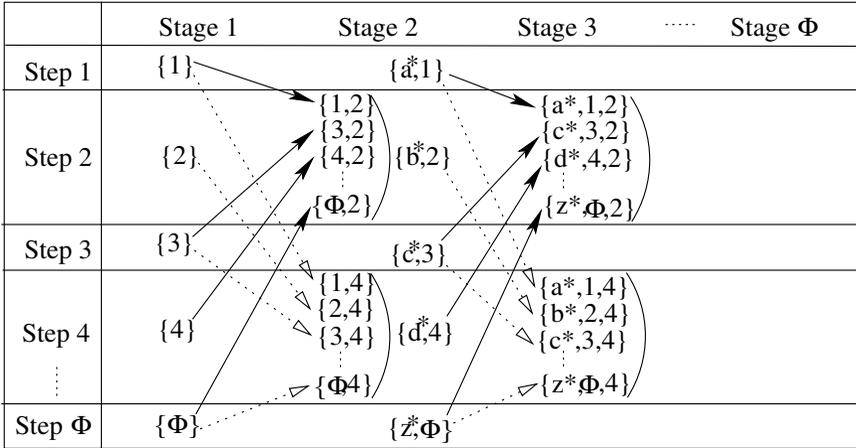


Fig. 14.9. Graphical depiction of the many-to-one traffic grooming heuristic. It is assumed that $a^* \neq 1, b^* \neq 2, c^* \neq 3, d^* \neq 4,$ and $z^* \neq \Phi$

The complexity of this heuristic was shown in [16] to be $O(\Phi^3 N^2 \sum_{c_k} |S_{c_k}|^2)$. It was also shown, through comparison to the exact approach, that this heuristic produces results which are between 10% and 38% of the exact solution.

14.4 Research Issues in Multipoint Traffic Grooming

Multipoint traffic grooming is a new field of research, and there are several open research problems in this field. As was evident from the discussion, the complexity of the optimal network design and session provisioning problem exceeds that of the normal unicast traffic grooming, which is known to be NP-hard. Therefore, approximate and heuristic approaches need to be developed for use with large networks, and increased traffic levels. In addition, session provisioning of multipoint traffic sessions under dynamic traffic conditions is an important problems, especially with the use of IP and MPLS over WDM. Tractable and efficient procedures are therefore required. The development of accurate performance models for the evaluation of blocking probabilities under this type of traffic is still an open area of research.

Finally, the survivable multipoint traffic grooming problem, where multipoint low-speed connections need to be protected, is still not receiving much attention in the literature. Recently, some work addressed the traffic grooming problem for survivable WDM networks; however only unicast traffic demands were considered. There is a need to propose and develop new approaches to provision and protect low-speed multipoint traffic demands.

References

1. B. Quinn and K. Almeroth, "IP multicast applications: Challenges and solutions." Internet Engineering Task Force (IETF) Request for Comments (RFC) 3170, Sept. 2001.
2. G. Rouskas, "Optical layer multicast: Rationale, building blocks, and challenges," *IEEE Network*, vol. 17, no. 1, pp. 60–65, 2003.
3. A. Hamad, T. Wu, A. E. Kamal, and A. K. Somani, "On multicasting in wavelength-routing networks," *Computer Networks*, vol. 14, pp. 3105–3164, Nov. 2006.
4. A. Hamad and A. E. Kamal, "A survey of multicasting protocols for broadcast-and-select single-hop networks," *IEEE Network*, vol. 16, pp. 36–48, July 2002.
5. G. Chowdhary and C. S. R. Murthy, "Grooming of multicast sessions in WDM mesh networks," in *Workshop on Traffic Grooming*, 2004.
6. D. Cavendish et al., "New transport services for next-generation sonet/sdh systems," *IEEE Communications*, vol. 40, pp. 80–87, May 2002.
7. R. Ul-Mustafa and A. E. Kamal, "Design and provisioning of WDM networks with multicast traffic grooming," *IEEE Journal on Selected Areas in Communications, Part II: Optical Communications and Networking*, vol. 24, pp. 37–53, Apr. 2006.
8. H. V. Madhyastha, G. V. Chowdhary, N. Srinivas, and C. Siva Ram Murthy, "Grooming of Multicast Sessions in Metropolitan WDM Ring Networks," *Computer Networks*, vol. 49, no. 4, pp. 561–579, 2005.
9. A. R. Billah, B. Wang, and A. A. Awwal, "Multicast traffic grooming in WDM optical mesh networks," in *Proceedings of the Conference on Global Communications (GLOBECOM)*, 2003.
10. A. L. Chiu and E. H. Modiano, "Traffic grooming algorithms for reducing electronic multiplexing costs in WDM ring networks," *IEEE Journal of Lightwave Technology*, vol. 18, pp. 2–12, Jan. 2000.
11. R. Dutta, S. Huang, and G. Rouskas, "On optimal traffic grooming in elemental network topologies," in *the Proceedings of Opticomm*, pp. 13–24, 2003.
12. G. Chowdhary and C. S. R. Murthy, "Dynamic multicast traffic engineering in WDM groomed mesh networks," in *Workshop on Traffic Grooming*, 2004.
13. A. Khalil, C. Assi, A. Hadjiantonis, G. Ellinas, and M. Ali, "On multicast traffic grooming in WDM networks," in *IEEE Int. Symp. on Computers and Communications*, pp. 282–287, 2004.
14. A. Khalil, A. Hadjiantonis, G. Ellinas, and M. Ali, "Sequential and hybrid grooming approaches for multicast traffic in WDM networks," in *Proceedings of the Conference on Global Communications (GLOBECOM)*, 2004.
15. E. Hernandez-Valencia, M. Scholten, and Z. Zhu, "The Generic Framing Procedure (GFP): An Overview," *IEEE Communications*, vol. 40, no. 5, pp. 63–71, May 2002.
16. R. Ul-Mustafa and A. E. Kamal, "Many-to-one Traffic Grooming in WDM Networks with Variable Aggregation Ratios," *IEEE Journal on Selected Areas in Communications, Part II: Optical Communications and Networking*, vol. 24, pp. 68–81, Aug. 2006.

Part III

Frontiers

Multi-Domain Traffic Grooming

Nasir Ghani, Qing Liu, Nageswara S. V. Rao, Ashwin Gumaste

15.1 Introduction

The last decade has seen many advances in high-speed networks. At the fiber level, dense wavelength division multiplexing (DWDM, Layer 1) has gained favor as a terabit solution, capable of “lightpath” circuit routing. Meanwhile next-generation SONET/SDH (NGS) (Layer 1.5) [1] has gained strong traction in the metro/edge, providing flexible “sub-rate” aggregation. Finally, Ethernet and IP networks (Layers 2, 3) have seen new quality of service (QoS) provisions via the differentiated services (Diff-Serv) and integrated services (Intserv) frameworks. Related control standards have also emerged, most notably multi-protocol label switching (MPLS) for Layer 2/3 flow-based QoS support. Further generalizations have even adapted this solution for “non-packet-switching” layers, i.e., generalized MPLS (GMPLS) [2].

As the above solutions undergo widespread deployment, a complex interconnection of different data-planes has resulted. From a theoretical aspect, this is embodied by an interconnection of multiple horizontal domains (i.e., networks) comprising of different vertical layers (i.e., technologies), as shown in Fig. 15.1 [3]. These segmentations are based upon various factors, such as technology, scalability, geographic, economic, administrative, etc. Moreover as the scale and reach of services expands, there is a pressing for “end-to-end” provisioning across heterogeneous domains, i.e., horizontal-vertical control. A good example is the growth in “e-science” applications [3]. However, inter-layer provisioning today is still done via manual provisioning of domain-specific control planes, giving high inefficiency and long lead times, i.e., hours to days.

Multi-domain networking has long been supported in IP and ATM networks. Moreover, new efforts are also introducing these capabilities in optical transport networks. Nevertheless, the broader topic of provisioning across heterogeneous network layers has not been addressed in detail. Note that there is a clear distributed multi-domain grooming aspect to this problem as different circuit/flow granularities are involved, i.e., gigabit wavelengths, “sub-rate”

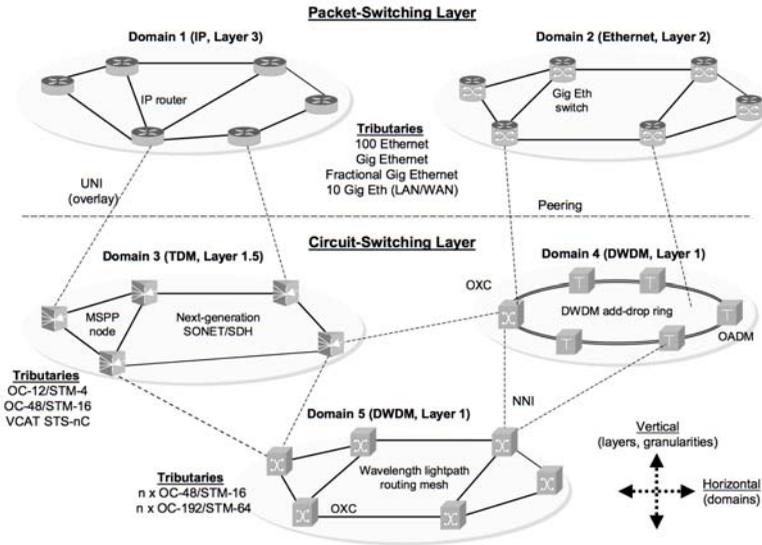


Fig. 15.1. Multi-layer, multi-domain networking infrastructures

SONET/SDH tributaries, MPLS label switched paths (LSP), etc. Overall, this is a very difficult problem owing to some key complicating factors. Foremost, there are no unified control plane standards for integrating multiple layers. In addition related provisioning algorithms are lacking, particularly those operating in distributed (intra, inter-carrier) settings with no/partial global state.

This chapter addresses the emerging area of multi-domain grooming and is organized as follows. First, Section 15.2 presents a brief survey of standards for multi-domain networks. Subsequently Section 15.3 surveys some work on multi-domain protocols and algorithms. Section 15.4 then delves into the relatively unexplored area of multi-domain grooming and details key directions in inter-layer routing, path computation, signaling, and survivability. Finally Section 5 presents overall conclusions and future directions.

15.2 Multi-Domain Standards

Over the years, a wide range of multi-domain networking standards have been developed by various standards organizations, including the IETF, ITU-T, and OIF. These efforts are briefly reviewed here.

15.2.1 ITU-T Standards

The ITU-T has matured a comprehensive automatically switched transport networks (ASTN) framework for multi-domain optical networks. The reference architecture for ASTN is G.8080 [2] and describes a group of components to control transport resources to setup, maintain, and release a client layer connections. However, the internal topology and connectivity of the underlying optical layers is not made visible to the client layer and is instead treated as a sub-network point pool (SNPP) link (“virtual” link). Now ASTN follows a hierarchical design in which designated entities perform control for single and multiple layers. For example, its routing hierarchy consists of areas with requirements for inter-area auto-discovery, auto-provisioning and auto-restoration. However, this model is flexible and each domain’s internal control plane can be tailored to the particular types and capabilities of the equipment within the domain. Nevertheless, this overall framework only addresses architectural design and no explicit protocols specifications are provided. It is here that developments within the IETF and OIF are of crucial importance.

15.2.2 OIF Standards

The OIF has developed several control standards for optical network interfacing, including a client–carrier user-network interface (UNI) and a carrier–carrier network node interface (NNI). The UNI implements bandwidth signaling for client devices (i.e., IP/MPLS routers) to request/release “optical” connections from underlying carrier SONET/SDH or DWDM domains, i.e., “optical dialtone”. Since there is no trust relationship here, carrier topology information is not propagated to the client side, i.e., overlay model [2]. The latest UNI 2.0 features much improved capabilities for security, bandwidth modification, etc. Meanwhile, the NNI interface implements inter-domain functions for reachability/resource exchange and setup signaling. Now two variants have been designed here, interior NNI (I-NNI) and external-NNI (E-NNI). The former interfaces nodes within the same administrative area whereas the latter serves adjacent areas. E-NNI relegates all interfacing issues to domain boundaries, thereby removing restrictions on domain-internal control and equipment interoperability. Furthermore, NNI adopts a link state routing approach, e.g., via single/stacked instances of GMPLS routing protocols.

The overall UNI–NNI combination facilitates rapid deployment of new services across multi-vendor “optical” domains. Moreover, this framework works for both optical DWDM and SONET/SDH domains, as facilitated by the different granularities supported in UNI/NNI signaling. Hence the NNI standard can support multi-layer interfacing at the circuit-switching level. However, since there is no resource exchange across the UNI, higher layer IP/MPLS networks must treat underlying “optical” links as tunneled links. This setup will be problematic for client layer routing since the number of Layer 3 adjacencies will grow per the square of border nodes, i.e., connection

mesh problem. Clearly, as the number of border nodes grows, this routing setup becomes less scalable, considering the amount of adjacency information has to be maintained. Hence clients may have to run data-plane inter-domain routing protocols across UNI-provisioned connection links.

15.2.3 IETF Standards

Well-established IP network architectures comprise a hierarchy of autonomous systems (AS) and routing areas (domains). Within areas, routers run associated interior gateway protocols (IGP) such as open shortest path first (OSPF) or intermediate-system to intermediate-system (IS-IS) [2]. These protocols use link-state routing to maintain link state databases (LSDB). Meanwhile at the inter-AS level, commensurate exterior gateway protocols (EGP) are used. The mainstay here is border gateway protocol (BGP) which runs between exterior gateway nodes and uses distance vector routing to provide end-point reachability exchange. In addition, OSPF also provides an additional level for disseminating “higher-layer” aggregated state between area border routers (ABR). Note that many operators will want to limit internal state dissemination in inter-carrier settings owing to privacy and security concerns.

With growing flow-level QoS requirements in IP networks, new extensions have also been proposed for OSPF. For example, OSPF-traffic engineering (OSPF-TE, RFC 2676) provides extensions and opaque link state (LSA) definitions for “QoS-related” link state to support advanced constraint-based routing (CBR). Meanwhile there have also been proposals for augmenting BGP messaging with added QoS attributes. For example, [4] exchanges QoS-enabled reachability information via new BGP message attributes. However, these offerings do not alter the inherent distance vector design. Finally, the hierarchical routing has also been implemented in older asynchronous transfer mode (ATM) networks via the private network-to-network interface (PNNI) protocol, i.e., via peer group leaders.

Meanwhile the GMPLS framework devises new routing/signaling versions to manage “circuit-switched” entities, e.g., TDM circuits and DWDM light-paths [2]. For example, new OSPF-TE opaque LSA definitions have been added for DWDM and SONET/SDH links. Hence TE databases (TEDB) can now store information fields for wavelengths/usages, timeslots/usages, shared risk link groups (SRLG) diversity, etc. New enhancements have also been added to the reservation protocol, RSVP-TE, to implement DWDM and SONET/SDH circuit setup/takedown, e.g., hard state, recovery, etc. Additionally, RSVP-TE also provides a loose route (LR) feature to specify a partial (high-level) node sequences along with signaling-based explicit route (ER) expansion to resolve the full path route. Finally a new link management protocol (LMP) has been developed for fault discovery and localization.

However, GMPLS does not provide any specific routing extensions for multi-domain BGP as its inherent distance-vector design is not well-suited

for circuit routing. Therefore, two-level OSPF-TE is the most germane inter-domain protocol for GMPLS. Nevertheless, OSPF-TE assumes a peer-to-peer model in which all nodes perform identical routing/signaling. Clearly, today's emerging diversified, multi-layer domains settings will prove problematic here. For example, many vendor solutions may not provide GMPLS support and instead use centralized vendor-proprietary network management systems (NMS) or operations support systems (OSS) and TL-1 messaging [1]. In these cases, proxy solutions can be used to translate external vendor-specific control formats [5].

The IETF is also addressing TE path computation via its path computation element (PCE) architecture [6, 7]. This framework allows path computation clients (PCC) to interact with PCE entities to resolve constraint-based routes. A PCE either resides in a standalone manner or is co-located with a network node and has access to domain-level resource/policy databases. This architecture provides a request/response protocol for PCC-PCE and PCE-PCE interaction, i.e., PCE protocol (PCEP) [7]. Finally, automated discovery allows PCC entities to locate a PCE with requisite capabilities, e.g., in terms of network layers, algorithms, backup path computation, etc. Two PCE computation models have also been proposed, centralized, and distributed [6]. In the former, all path computation is performed by a single PCE entity. Although this is realistic for intra-domain scenarios, the broader multi-domain case necessitates this PCE to have global resource and policy visibility. Clearly, a single entity also poses notable scalability and reliability limitations.

Meanwhile, distributed computation uses multiple PCE entities to resolve end-to-end paths. This approach also supports "domain-specific" constraints, such as bandwidth/delay in IP domains and wavelength/color continuity in DWDM networks. Moreover, distributed PCE computation is very germane for multi-domain/multi-layer/multi-carrier settings with partial visibility. Herein, two approaches are outlined to handle the contingencies of varying levels of "global" state, i.e., multi-PCE path computation with and without inter-PCE signaling [7]. Overall both of these schemes induce close couplings between inter-domain routing and signaling.

Multi-PCE computation without inter-PCE signaling requires the head-end PCE (i.e., signaled by PCC) to compute a partial or loose route to the destination. This route can be a "domain-level" sequence of border gateways which is then inserted into appropriate downstream signaling messages. Here receiving entities (border gateways) consult their own PCE entities to "expand" the strict-hop sequences across their domains to the next partial/loose hop. This "inter-PCE" approach fits in well with the above-mentioned RSVP-TE LR/ER features. Meanwhile, multi-PCE computation with inter-PCE signaling requires the head-end PCE to recursively request other PCE entities to compute sub-path segments. This setup usually returns a complete strict-hop source-destination route to the requesting PCC and is more suitable for little/no inter-domain visibility. Alternatively, in inter-

carrier settings, this approach may still return a loose route sequence, thereby preserving intra-domain route confidentiality, see [7].

15.3 Research Studies

Many multi-domain research studies have been conducted. However, most of these efforts have focused on single-layer (homogeneous) packet/cell-switching networks, Layers 2–3. Recently some work on multi-domain optical networks (Layer 1) has also emerged. As yet, very few studies have looked at multi-domain/multi-layer inter-networking.

15.3.1 Multi-Domain Packet/Cell-Switching Networks

Many studies have looked at multi-domain ATM (Layer 2) and IP (Layer 3) networks. In particular, graph-based topology abstraction schemes for state reduction [8–11] have been a key focus area. These strategies transform a domain topology into a reduced “virtual” graph with fewer abstract vertices and edges. This is typically done by a designated domain entity, e.g., routing area leader (RAL) [2], which then propagates the abstract link state to other gateway nodes to maintain “globalized” state. Expectedly, the computational entity must have access to intra-domain state. Overall, topology abstraction allows operators to control internal state dissemination and effectively couples intra- and inter-domain routing.

An early study of abstraction in hierarchical ATM PNNI networks is presented in [8]. This effort details algorithms for computing complex node representations to minimize the number of edges/links in a peer group summarization. The findings here show much-reduced computation times and over an order magnitude reduction in domain state. Also, [9] extends abstraction to multi-domain IP QoS networks, summarizing bandwidth state using various topologies, i.e., star, mesh, tree, and spanner graphs. These reductions are tested with various path computation strategies, including widest shortest, shortest-distance, etc. The overall findings confirm improvements in routing scalability and reduced routing fluctuation. Meanwhile, [10] looks at topology aggregation in directed graphs with additive link metrics (delay, cost). This problem is treated from an information-theoretic perspective and related bounds for compression distortion are derived based upon the asymmetry constant. Furthermore, [11] develops schemes to incorporate both bandwidth and delay parameters via line segmentation techniques. Various Dijkstra-based heuristics are then used to compute paths meeting dual bandwidth and delay constraints. Overall, the results show improved success rates and lower crankback messaging loads.

Finally, studies have also looked at QoS enhancements for BGP distance vector routing. For example, [4] proposes messaging extensions to convey

additional QoS state for paths to destination domains/prefixes. This is coupled with pre-engineered QoS-based service level specifications (SLS) between providers along with “meta-QoS class” planes extending across multi-carrier AS. Results for sample Internet topologies show best performance when paths selection is done using both bandwidth and delay constraints. In addition, messaging overheads are shown to be minimal.

15.3.2 Multi-Domain Optical Networks

Multi-domain DWDM networks are also becoming a key focus. To date, most efforts in this area have focused on distributed lightpath RWA and signaling in all-optical and opto-electronic networks. For example, [12] details a domain-by-domain scheme in which gateways maintain complete (alternate) route state across all-optical and opto-electronic networks. Detailed simulation shows the overall effectiveness of the proposed model. However, this setup is more favorable to BGP-type implementations and related resource propagation (path dissemination) results are not shown.

Meanwhile, [13] studies DWDM networks comprising of multiple segments, i.e., domains. Here, three different inter-domain wavelength routing algorithms are developed, i.e., end-to-end (E2E), concatenated shortest path (CSR), and hierarchical routing (HIR). The E2E basically assumes a “flat” globalized graph whereas HIR assumes a hierarchical graph with segments summarized as nodes. Meanwhile CSR simply uses local information to perform segment-by-segment routing. Numerical results with a specialized mesh-torus topology show significant blocking reduction with the E2E scheme for “multi-segment” requests. However, the CSR scheme does not provide any intra-domain state and further inter-domain routing algorithms are not addressed.

Now, very few studies have looked at routing in multi-domain DWDM networks. Here, it is generally very difficult to re-apply IP/MPLS abstraction schemes since resource constraints are different. Hence recent studies have looked at new DWDM-based abstraction schemes. For example, [14] presents a theoretical treatment of abstraction with border node wavelength conversion. Here, various information models are developed and lightpath selection is modeled as a Bayesian decision problem. The findings for bus topologies show that scalable information models can achieve a good trade-off between performance loss and the amount of network state. Nevertheless, inter-domain routing and RWA issues are not addressed.

Meanwhile, [15] tables a hierarchical inter-domain solution using simple-node abstraction. Nevertheless, signaling provisions are not considered here and only the all-optical case is handled. Conversely, [16] presents a much more detailed study of abstraction in multi-domain all-optical/opto-electronic DWDM networks. Namely, graph theoretic RWA schemes (via k-shortest path heuristics) are used to generate full mesh and star abstractions. Furthermore, detailed inter-domain link-state routing protocols and distributed lightpath

RWA algorithms are also specified (using OSPF-TE, RSVP-TE). Findings show very good blocking reduction and lower inter-domain signaling loads.

Dynamic sub-path protection for DWDM networks has also been considered. For example, [17] segments end-to-end lightpaths into independent “protection domains” and implements shared segment protection with/without wavelength conversion. The algorithm pre-defines self-healing cycles in the network topology and attempts to allocate self-healing loops for lightpaths accordingly. Overall, it is shown that maximum sharing can be achieved by properly assigning wavelengths to the shared channels. Nevertheless, these schemes assume full global state, i.e., “flat/single-domain” network, and hence their application in distributed multi-layer settings is not straightforward. Finally, Layer 1 virtual private networks (Layer 1 VPN) provisioning in multi-domain SONET/SDH network has also been studied in [18]. Layer 1 VPN represents a new service model in which transport-layer resources are partitioned to build multiple “virtual infrastructures” over a single physical network. Here, the authors develop a novel OSPF-TE link model to share time-slots between client VPN topologies. Furthermore, non-adjacent L1 VPN clients are interconnected via “virtual link” multi-domain SONET/SDH connections. Overall, results show good sharing efficiency and blocking reductions with shortest-widest path selection. However, all virtual links here are pre-specified (static), as per end-user demands, and inter-domain routing is not done.

15.3.3 Multi-Domain/Multi-Layer Networks

As mentioned in Section 15.1, there is a clear grooming aspect to multi-domain/multi-layer networks. For example, multiple fine-granularity label switched paths (LSP) originating in Layer 3 domains can be aggregated over coarser DWDM lightpaths. In a similar manner, sub-rate SONET/SDH circuits can also be groomed over DWDM lightpaths. Now, even though grooming algorithms have been widely studied over the years, most studies have focused on idealized settings with full resource state knowledge across all layers. The broader topic of distributed multi-domain grooming, i.e., in the presence of limited global state, is largely unexplored and only a handful of studies have been conducted to date.

In particular, [19] presents one of the first studies on distributed SONET-DWDM provisioning in multi-domain networks. Here, the authors define a multi-segment graph model (with boundary grooming) and evaluate various path selection schemes, e.g., centralized (full-knowledge), domain-by-domain (local knowledge), and hierarchical source routing (partial inter-domain knowledge). The latter approach only propagates domain-internal state for a specified granularity levels, although detailed specifications for state compression are not given. Overall, results show much-improved blocking performance with increasing levels of inter-domain state. Furthermore, [20] studies grooming in SONET-DWDM networks and uses threshold-based triggers to add/delete lightpaths. Here, both centralized (global) and distributed schemes

are developed, with the latter only using traversing path state at a node, i.e., no routing. Overall, results show good performances, contingent to grooming ports and request granularities. However, no inter-domain considerations are discussed.

Finally, [21] tables a novel framework for direct Ethernet-over-DWDM integration. This scheme devises an integrated signaling-based approach to simultaneously provision full wavelength and fine granularity Ethernet virtual connections (EVC) at the optical layer. In contrast to conventional provisioning, signaling messages for sub-lambda EVC requests are only processed at the source node of the lightpath. Nevertheless, the authors assume that the underlying DWDM network is relatively static and hence routing is only done at the Ethernet layer. Detailed performance results are presented for various routing update strategies using crankback. Overall, results show minimal increase in setup latencies and signaling loads.

15.4 Open Problems in Multi-domain Grooming

Despite some standards progress, the study of multi-domain/multi-layer grooming remains in its infancy. Now, it is clear that the scope/scale of the problem necessitates distributed, decentralized solutions here. Along these lines, some key open problems need to be addressed, i.e., multi-domain routing, constraint-based path computation/signaling, and survivability.

15.4.1 Multi-layer Routing and State Exchange

Multi-domain routing is a very challenging area, especially in the presence of multiple layers. The most notable issue is how to propagate different types of link or domain state in a such way that a scalable and secure routing framework can be achieved. By and large, multi-domain grooming will benefit from the availability of “global” inter-domain state. Moreover, as per Sections 15.2.3 and 15.3.1, link-state routing is the most suitable scheme for constraint-based provisioning of QoS LSP or TDM/DWDM circuit paths. Hence, it is very desirable to implement some form of hierarchical routing at select border gateways to disseminate various types of link state (as shown in Fig. 15.2):

- **Physical inter-domain links:** These links interconnect nodes in different domains and can span a full range of types, e.g., Gigabit/10 Gigabit Ethernet, SONET/SDH OC-n links, DWDM links, etc. Hence, related multi-layer TEDB must maintain state for all of these link types either via running multiple routing protocol instances or by using a single unified routing protocol. For example, two-level OSPF-TE already provides LSA definitions for both packet-switching links and “non-packet” TDM and DWDM links.

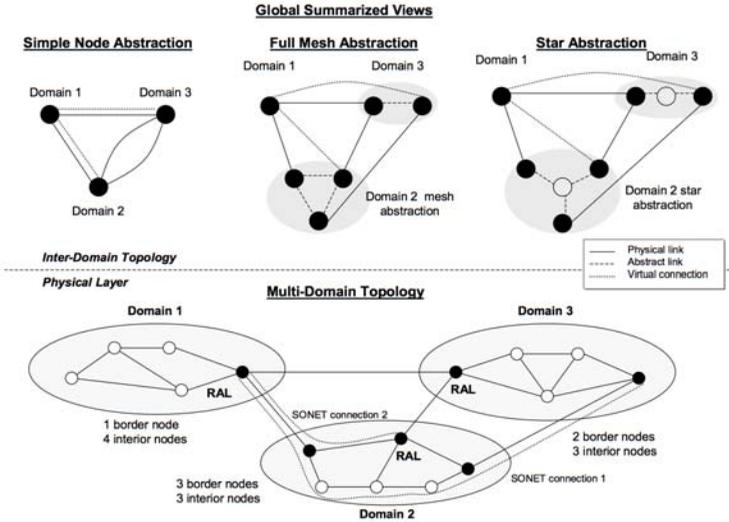


Fig. 15.2. Topology abstraction and link types

- **Abstract links:** Abstract links are computed entities used to summarize “layer-specific” domain-level resources and do not reflect physical elements [9]. Although abstraction has been well-studied for IP networks (Section 3.1), circuit-switching networks are more specialized, see initial work in [16]. Hence, further considerations are needed for SONET/SDH grooming links, VCAT features, inverse multiplexing, etc. In addition, new abstractions must capture survivability-related information via abstract links, e.g., fiber/span protection, diversity/risk groups, shared/dedicated resources, etc. In general, expanded attributes can also be developed as these link types need not be constrained to physical link types (in respective domains).
- **Virtual link connection:** In general, lower level SONET/SDH and DWDM connections/sub-connections will appear as virtual links to higher-layer devices (i.e., IP/MPLS, Ethernet nodes). These entities can thus be treated as logical TE link (adjacencies) between the two end-points, see Fig. 15.2. Therefore, it will be beneficial to also advertise the resource levels for these entities at the inter-domain level as well, since doing so can yield very good gains in grooming efficiencies. However, there is a clear scalability problem here as the number of such connections can grow in the order of the square of the number of border nodes, i.e., the connection mesh problem as discussed in Section 15.2.2. Hence, a key challenge here will be to develop link aggregation techniques to effectively compress virtual connection link state.

In light of the above, it is also important to consider associated update triggering strategies. To date, various policies have been developed here to achieve update frequency reduction/scalability, e.g., timer-based, absolute change, relative change, and hysteresis-based [20]. All schemes here—except for the former type—are threshold-based. Extensive results for single-domain networks show that relative updates are generally most effective in lowering routing loads and ensuring rapid change sensitivity [20]. Now physical inter-domain link updates can be treated in much the same manner as intra-domain links. Namely, if the underlying resources (bandwidth, timeslots, wavelengths) change as per the chosen metric, an update can be flooded by the sourcing gateway. However, careful analysis is needed to delimit related inter-domain holdoff timers (IHT), which will tend to be larger than corresponding intra-domain values. Virtual link connections can also be treated as such since they are directly associated with underlying connections. Nevertheless, since the number of such connections can be large, further research into scalable triggering policies is needed. Meanwhile, triggering policies for abstract links are more complicated as they pertain to “non-existent” entities. Here, various policies have been developed for abstract IP and DWDM links, see [9, 16], and these can be adapted further. For example, a basic approach is to compute abstractions on a fixed time interval. However, this scheme yields excessive messaging/computational loads for small update intervals and inaccurate state for large update intervals. It is here that relative change measures can give much better performance. For example, [16] uses periodic abstractions but only sends updates if sufficient relative change occurs.

15.4.2 Multi-Layer Grooming and Signaling

Path computation implements the core of the multi-domain grooming function. As opposed to single-domain settings, full topology information in multi-domain networks is generally lacking and path computation has to be done using aggregated (i.e., partial, dated) topology and resource state. Moreover, the existence of multiple layers implies an inherent grooming component, again differing notably from single domain settings which tend to operate at the same layer, e.g., either packet, timeslot, wavelength. Overall, these aspects make multi-layer grooming involving a very challenging problem. In the end, however, the goal here is to resolve and set up full end-to-end routes in accordance with desired constraints and leverage existing standards, particularly GMPLS. Now owing to the distributed nature of the problem it is evident that setup signaling will be required. Although many specific schemes can be envisioned here, few have been analyzed in detail. Hence two broad strategies are outlined here, i.e., hierarchical and per-domain.

Hierarchical Grooming

Hierarchical grooming uses the inter-domain TEDB state to compute “skeleton” routes from the source domain to the destination domain. For

most practical purposes, this path constitutes the end-to-end border node sequence. This information is then coupled with end-to-end signaling mechanisms (RSVP-TE) to expand the full node sequence, i.e., multi-PCE path computation without inter-PCE signaling [7]. Alternatively, this information can also be coupled with PCEP signaling to fill in an exact node sequence, i.e., multi-PCE path computation with inter-PCE signaling [7].

Now consider the actual computation of “skeleton” routes. This is a very complex problem owing to the multiple resource layers involved and the presence of partial/abstracted inter-domain state. Hence, two key areas need detailed investigation here. Foremost, the computing entity (e.g., source domain PCE) must adequately “transform” multi-layer TEDB state to build a “global” multi-layer (granularity) view of the network. The most expedient approach here is to use graph transformations to summarize all link types—physical, abstract, and virtual (Section 15.4.1). This topic is largely unexplored, although augmented graph models have been developed in [9] for multi-layer SONET-DWDM grooming networks (full global state assumption). Clearly, new schemes are needed for adding edges/vertices to capture virtual link connections as well. Next, modified shortest-path schemes need to be studied to compute constrained “skeleton” paths over these transformed graphs. It is here that the grooming aspect can be considered, particularly when “higher-layer” sub-rate connections are to be routed over lower-layer circuits (virtual link connections).

Most regular (non-multi-domain) grooming studies use a “two-step” computation approach. Namely, setup is first attempted over existing higher-layer “virtual” links and pending failure, re-attempted at lower transport layers. Overall, this approach yields very good resource efficiency/packing and is equally relevant in inter-domain scenarios. For example, the initial attempt can search for a “skeleton” route on only those links in the augmented graph (physical inter-domain, abstract, virtual connection) which run at the same granularity/layer as the sourcing node. Here various graph theoretic schemes can be developed to achieve trade-offs between objectives such as hop count, load balancing, minimum cost, e.g., see studies on widest-shortest, shortest-widest, and minimum cost [9].

If a feasible “skeleton” path is not found, further attempts can try to set up new underlying virtual link connection between one/more border gateways. However, this step opens up a whole new dimension that is not present in single-layer multi-domain settings. Namely, the main challenge is to select the actual border node pair(s) between which to initiate virtual connection link(s). This concern has not been addressed in most grooming studies which exclusively assume that all higher layer nodes are directly connected to underlying transport nodes. Clearly this is not the case in general settings in which interior IP/MPLS nodes will not be directly connected to lower-layer SONET/SDH or DWDM transport nodes. One possibility is to use exhaustive search algorithms to enumerate multiple virtual links combinations, e.g., as

done in [16]. Another simpler possibility may be to pre-engineer static virtual connection links using offline optimizations.

Upon computation of a feasible “skeleton” path, appropriate setup signaling schemes are needed as well. In general, these schemes can leverage the ubiquitous RSVP-TE protocol. Specifically, resources can be reserved on physical inter-domain and “tunneled” virtual connection links and path sequences “expanded” for loose route segments. With regards to the latter, border gateways receiving a setup request will basically run intra-domain algorithms to route requests across their domains. Now if the incoming request is of the same granularity, the border node can compute (or query PCE to compute) a regular traversing as per path constraints. However, in the more general case where the border node represents a lower layer, further parameter translation and intra-domain grooming will be needed, and a wide range of existing (intra-domain) grooming schemes can be re-applied. Furthermore, new RSVP-TE proposals are also tabling LSP “stitching” features to inter-connect diverse end-to-end “segments” (layer paths) into a single connection. Again, these can be leveraged as well.

Overall, hierarchical computation can achieve some level of path “optimality” across domains, i.e., contingent to level of inter-domain state. An “optimal” path here is defined as that computed in the absence of domain partitioning, i.e., for “flat” topology [7]. However, detailed studies are really needed to assess the impact of abstracted state on such “optimality”.

Per-Domain Grooming

Per-domain distributed grooming is ideal for settings with minimal inter-domain visibility. Namely, end-to-end routes are expanded in a domain-by-domain manner with each domain specifying its “next-hop” along with the complete internal route via an egress gateway. Here, intra-domain path expansion will be largely similar to the above-detailed hierarchical approach. However at the inter-domain level, the main challenge is how to specify the next domain. If no inter-domain state is available, the only solution may be to make fixed choices, i.e., pre-specified or by consulting BGP tables [7]. Alternatively, it may be much more beneficial to incorporate “global” state to dynamically the downstream domains, if such state is available. Nevertheless, the detailed evaluation of these schemes in mixed-layer networks has not been conducted. In general, per-domain path computation (without inter-domain state) will suffer from higher blocking and lower resource efficiencies as compared to hierarchical computation. Hence, resultant routes will likely be feasible and opposed to “optimal”.

Note that per-domain grooming performance can be improved by leveraging inter-domain crankback signaling. Although [22] studies this approach in SONET-DWDM networks, more investigations are needed. Specifically, commensurate schemes can be designed using intermediate domain or full end-to-end crankback strategies. Furthermore, it may be beneficial to also include active path state in crankback messages in order to improve the

re-try process. Although multi-layer crankback can yield improved resource efficiencies, it poses key trade-offs in inter-layer control plane complexity. Namely, this solution may lower inter-layer routing overheads (e.g., versus hierarchical computation) but at the same time will likely increase inter-layer signaling and setup latencies. These issues need to be closely studied in order to determine the most amenable grooming strategies, i.e., as per connection granularities, arrival rates, holding times, etc.

15.4.3 Survivability

With limited or no visibility into domain-level state, setting up survivable connections is an extremely complicated issue. Hence, the simple approach of computing diverse routes by source nodes (with partial global state) can only guarantee diversity at the physical inter-domain link level. Clearly, further provisions are needed to avoid link overlap during subsequent intra-domain explicit route expansion even with disjoint abstract links computed. Overall, survivable grooming is a relatively new topic area. For example, [23] has recently proposed some protection schemes for SONET-DWDM grooming networks, e.g., protection-at-lightpath (PAL) and protection-at-connection (PAC). Nevertheless, extending these concepts into multi-domain networks lacking full resource diversity state is very challenging. Furthermore, it is also very desirable to extend some form of multi-tiered survivability support as well, e.g., dedicated, shared, non-protected, etc. To address this topic, two survivability strategies can be considered, protection and restoration.

Protection schemes pre-compute diverse primary/backup routes and are well suited for stringent services. Now in the context of multi-domain/multi-layer networks, a simple approach can be to apply SONET strategies in which dual/multi-homed inter-domain links are coupled with robust intra-domain protection schemes. However, these setups are very costly and resource inefficient. Hence, more elaborate end-to-end domain disjoint techniques can be devised, which are generally more amenable to hierarchical grooming with inter-domain state visibility. For example, loose routing algorithms can be extended to compute domain disjoint primary/back “skeleton” paths over augmented inter-layer graphs. Associated setup (loose route expansion) sequences can then be signaled in parallel. Alternatively, the domain-disjoint requirement can be relaxed to allow sharing of abstract links between primary/backup routes. However, the latter approach mandates complicated signaling provisions to ensure intra-domain disjointness in common domains. Again, many of these concerns are open issues and the combined impact of intra/inter-domain schemes on end-to-end recovery still needs to be addressed.

Conversely protection can also be incorporated with the per-domain grooming approach of Section 15.4.2. For example, primary/backup paths can be signaled in a two-step sequential manner where the working route is used to avoid backup link overlaps. Obviously, this approach will have higher setup delays and will be susceptible to “trap” topologies [7]. To address these

concerns, modified crankback schemes can be considered to achieve protection path diversity with parallel signaling, e.g., as tabled in RSVP-TE extensions for record, exclude and associated route objects (RRO, XRO, ARO) [24]. To date, however, most of these concepts have only appeared various standards draft submissions (MPLS-level only), and more detailed analyses and refinements are indeed lacking.

Finally, multi-domain restoration can also be done using post-fault crank-back recovery and this is most germane for connections with lower latency/guarantee stringencies. Along these lines, various signaling provisions need to be devised, e.g., such as proposed RSVP-TE XRO extensions. However, further studies are needed to analyze the trade-offs between resource efficiencies and signaling loads and also compare versus pre-provisioned protection.

15.5 Conclusion

This chapter studies multi-domain grooming in heterogeneous layered networks, a challenging and largely unexplored problem. Along these lines related standards and research studies are surveyed and key open research problems identified. In particular, the latter issues pertain to multi-layer/multi-domain routing, path computation and signaling, and survivability. Finally, future studies can delve into even more advanced multi-domain problems, e.g., VPN provisioning, advance reservation, etc.

References

1. N. Ghani, et al., "Metropolitan Optical Networks," *Optical Fiber Telecommunications IV*, Academic Press, pp. 329–403, March 2002.
2. G. Bernstein, B. Rajagopalan, D. Saha, *Optical Network Control-Architecture, Protocols and Standards*, Addison Wesley, Boston 2003.
3. N. Rao, et al., "Ultra Science Net: Network Testbed for Large-Scale Science Applications," *IEEE Communications Magazine*, Vol. 3, No. 4, pp. S12–S17, November 2005.
4. D. Griffin, et al., "Interdomain Routing Through QoS-Class Planes," *IEEE Communications Magazine*, Vol. 45, No. 2, pp. 88–95, February 2007.
5. T. Lehman, et al., "Control Plane Architecture and Design Considerations for Multi-Service, Multi-Layer, Multi-Domain Hybrid Networks," *IEEE INFOCOM 2007 HSN Workshop*, Anchorage, Alaska, May 2007.
6. A. Farrell, "A Framework for Inter-Domain Traffic Engineering," *IETF Re-quest for Comments (RFC) 4726*, November 2006.
7. A. Farrell, J. Vasseur, J. Ash, "A Path Computation Element (PCE)-Based Architecture," *IETF Request for Comments (RFC) 4655*, August 2006.
8. I. Iliadis, "Optimal PNNI Complex Node Representations for Restrictive Costs and Minimal Path Computation Time," *IEEE/ACM Transactions on Network-ing*, Vol. 8, No. 4, August 2000.

9. F. Hao, E. Zegura, "On Scalable QoS Routing: Performance Evaluation of Topology Aggregation," *IEEE INFOCOM 2000*, pp. 147–156, 2000.
10. B. Awerbuch, Y. Shavitt, "Topology Aggregation for Directed Graphs," *IEEE/ACM Trans. on Networking*, Vol. 9, No. 2, pp. 82–90, February 2001.
11. K. Liu, K. Nahrstedt, S. Chen, "Routing with Topology Abstraction in Delay-Bandwidth Sensitive Networks," *IEEE/ACM Transactions on Networking*, Vol. 12, No. 1, pp. 17–29, February 2004.
12. X. Yang, B. Ramamurthy, "Inter-Domain Dynamic Routing in Multi-Layer Optical Transport Networks," *IEEE GLOBECOM 2003*, San Francisco, CA, December 2003.
13. Y. Zhu, A. Jukan, M. Ammar, "Multi-Segment Wavelength Routing in Large-Scale Optical Networks," *IEEE ICC 2003*, Anchorage, AL, May 2003.
14. G. Liu, C. Ji, V. W. S. Chan, "On the Scalability of Network Management In-formation for Inter-Domain Light-Path Assessment," *IEEE/ACM Transactions on Networking*, Vol. 13, No. 1, pp. 160–172, January 2005.
15. S. Sanchez-Lopez, et al., "A Hierarchical Routing Approach for GMPLS-Based Control Plane for ASON," *IEEE ICC 2005*, Seoul, Korea, June 2005.
16. Q. Liu, et al., "Hierarchical Inter-Domain Routing and Lightpath Provisioning in Optical Networks," *OSA Journal of Optical Networking*, Vol. 5, No. 10, pp. 764–774, October 2006.
17. P. Ho, H. T. Mouftah, "A Novel Survivable Routing Algorithm for Segment Shared Protection in Mesh WDM Networks with Partial Wavelength Conversion," *IEEE JSAC*, Vol. 22, No. 8, pp. 1539–1548, October 2004.
18. N. Ghani, et al., "Layer 1 VPN Services in Distributed Next-Generation SONET/SDH Networks With Inverse Multiplexing," *OSA Journal of Optical Networking*, Vol. 5, No. 5, pp. 367–382, May 2006.
19. Y. Zhu, et al., "End-to-End Service Provisioning in Multi-Granularity Multi-Domain Optical Networks," *IEEE ICC 2004*, Paris, France, June 2004.
20. R. Alnuweri, et al., "Performance of New Link State Advertisement Mechanisms in Routing Protocols with Traffic Engineering Extensions," *IEEE Comm. Magazine*, Vol. 42, No. 5, pp. 151–162, May 2004.
21. A. Hadjiantonis, et al., "Evolution to a Converged Layer 1, 2 in a Global-Scale, Native Ethernet Over WDM-Based Optical Networking Architecture," to appear in *IEEE Journal of Selected Areas in Communications*, 2007.
22. I. Widjaja, et al., "A New Approach for Automatic Grooming of SONET Circuits To Optical Express Links," *IEEE International Conference on Communications (ICC) 2003*, Anchorage, AL, May 2003.
23. C. Ou, et al., "Traffic Grooming for Survivable WDM Networks-Shared Protection," *IEEE JSAC*, Vol. 21, No. 9, pp. 1367–1383, November 2003.
24. C. Lee, et al., "Exclude Routes-Extensions to RSVP-TE," *IEEE Communications Magazine*, November 2006.

Grooming of Scheduled Demands in Multi-Layer Optical Networks

Maurice Gagnaire, Josué Kuri*, Elias A. Doumith

16.1 Introduction

Introducing multiple switching granularities in transport networks aims in part at reducing their cost by taking advantage of traffic aggregation. As a matter of fact, an appropriate aggregation of low-rate connections into high-rate connections¹ and the use for the latter of a switching technology whose cost-per-bit is lower than for the former can result in lower infrastructure costs. The economic gain, when compared to single switching granularity networks, depends on both traffic-related factors such as the time/space distribution of connections and equipment or network architecture factors such as the cost of ports and matrices for each switching granularity, the topology of the network, the provisioning algorithms used for routing and grooming of connections, etc.

In other chapters of this book, traffic grooming has been investigated in contexts such as protection/restoration, multicast, and multi-domain networks. In this chapter, we consider the problem of efficiently grooming (a) electrical “subwavelength” demands into lightpaths, and (b) lightpaths into wavebands. Whereas most investigations deal with either permanent or random traffic demands, we focus in this chapter on a particular class of traffic demands referred to as scheduled demands, or SxDs, introduced in [1, 2]. A scheduled traffic demand corresponds to a connection or set of connections for which the network operator knows in advance the set-up and tear-down dates. These connections could be requested, for example, by a company to connect its headquarters to its production centers during office hours, or to interconnect its data centers during the night, when database backups are performed. Since companies tend to have well-established operations and procedures, they can provide relatively accurate information to the network operator

* This work was carried out while the author was at ENST Paris. The opinions expressed in this paper are those of the author and not necessarily those of Orange Labs.

¹ For example, Low Order (LO) and High Order (HO) SONET/SDH connections.

about when the bandwidth is really needed. An operator can leverage the knowledge about the timing of demands to make a more effective use of its network resources.² For example, a same set of switch ports and transponders can be allocated to multiple connections routed over a same path if the operator knows in advance that the connections are time-disjoint. Conceptually, a permanent connection is a special case of a scheduled demand for which the set-up and tear-down dates are $-\infty$ and ∞ , respectively. In practice, the set-up and tear-down dates of a permanent connection correspond to the beginning and end dates of the planning period considered in a network design exercise. Therefore, by solving the problem of provisioning scheduled demands, we also solve the problem of provisioning permanent demands.

From a mathematical modeling point of view, an appealing characteristic of scheduled demands is that they capture the dynamic changes of the traffic load in a *deterministic* manner, which eases the use of well-known optimisation techniques such as heuristics [3, 4, 5], meta-heuristics [6], linear programming [7, 8, 9], etc.

A scheduled demand is formally characterized by a tuple (s, d, n, α, β) in which s and d represent the source and destination nodes, n is the requested capacity, and α and β represent the set-up and tear-down dates. Two forms of scheduled demands are considered in this chapter: Scheduled Lightpath Demands (SLDs) and Scheduled Electrical Demands (SEDs), also known as “subwavelength” demands. For SLDs, the capacity n is expressed in number of lightpaths (each one with a nominal rate of either 2.5 or 10 Gbps), and for SEDs, n is expressed as a fraction of a lightpath’s capacity. For example, $n = 0.4$ corresponds to 1 Gbps with 2.5 Gbps lightpaths.

For the problem of grooming SEDs into SLDs, we develop in this chapter a mathematical formulation to quantify the number of Electrical Cross-Connect (EXC) and Wavelength Cross-Connect (WXC) ports required to provision a set of SEDs in a given topology [10, 11, 12, 13]. For the problem of grooming SLDs into wavebands, we develop a different formulation which quantifies the monetary cost of provisioning a set of SLDs in a network in which an integrated Wavelength Cross-Connect (WXC) / Waveband Cross-Connect (BXC) equipment is present in each node.³ We refer to the former as the electrical grooming problem and to the latter as the optical grooming problem.

² This is analogous to what airlines do: passengers that buy tickets well in advance are rewarded with lower fares because early bookings reduce the uncertainty of the future demand, which allows the airline to make a more effective use of its resources.

³ An EXC is essentially the same equipment that has been referred to as a DXC in earlier chapters, and WXC and BXC are both OXCs, operating at different wavelength selectivities. - Editors

16.2 Architecture of Network Nodes

Figure 16.1 depicts the architecture of a multi-granularity switching cross-connect (MG-OXC) that integrates an electrical cross-connect (EXC), a wavelength cross-connect (WXC), and a waveband cross-connect (BXC). The composite signal received on the input fibers of the BXC is demultiplexed to extract individual wavelengths which are in turn multiplexed into wavebands that arrive at the input ports of the BXC switching fabric. The BXC cross-connects input wavebands to either output or drop ports. Dropped wavebands are demultiplexed to extract individual wavelengths which are cross-connected by the WXC to either output or drop ports. Dropped wavelength connections are terminated by the EXC, which extracts “subwavelength” connections (or flows) that are sent to client equipment (e.g., routers) through the EXC drop ports. The MG-OXC has the capability of dropping full lightpaths directly from the WXC to client equipment. In the opposite direction, “subwavelength” connections (or flows) received from client equipment through the EXC add ports are groomed into lightpaths, which are received by the WXC through its add ports and groomed into wavebands that are received by the BXC through its add ports and groomed into fibers. The WXC of the MG-OXC has the capability of adding lightpaths directly from client equipment.

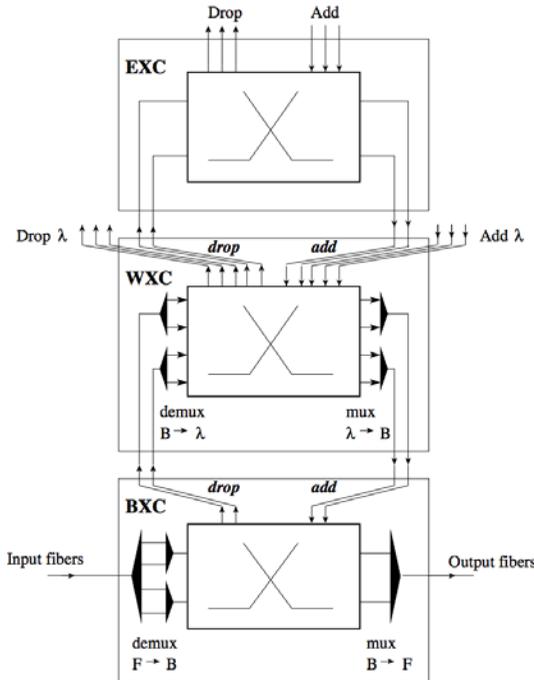


Fig. 16.1. Multilayer node architecture

In this chapter, we deal with electrical and optical grooming as two separate problems. In the former case, the network consists of EXC/WXC nodes (no BXC), whereas in the latter, the network consists of WXC/BXC nodes (no EXCs).

16.3 Rationale of Grooming Strategies

The goal of grooming strategies in multi-layer networks is to reduce the overall management burden and network cost (compared to single-layer networks) by efficiently grouping a large number of low-rate connections into a relatively small number of high-rate connections. The overall network cost can be reduced if the cost-per-bit is lower in the network layer bearing the high-rate connections than in the upper layers, and if the cost of grouping/ungrouping connections does not outweigh the savings of the high-rate connections' layer. Architectures like the one of Fig. 16.1 may reduce the network cost if BXC are less expensive than WXC, and WXC are in turn less expensive than EXC. This tends to be the case with state-of-the-art technology since EXCs require electronic switching fabrics and opto-electronic converters on I/O ports, whereas WXC and BXC require potentially less opto-electronic devices and can be based on less expensive optical switching fabrics.

The presence of transit traffic in a multi-granularity switching cross-connect can reduce the number of required EXC and WXC ports since wavebands bearing transit lightpaths are switched directly from input to output ports of the BXC without requiring any port on the WXC. Likewise, lightpaths bearing transit "sub-wavelength" traffic are switched directly from input to output ports of the WXC without requiring any port on the EXC. Lightpaths on the input ports of the WXC are either recombined into wavebands through the output ports or sent to either client equipment or input ports of the EXC. The EXC extracts "sub-wavelength" connections (or flows) which are sent to client equipment through the drop ports or recombined into new lightpaths sent to the WXC through the output ports.

Provisioning in a multi-layer network involves essentially the definition of both, a path for each SEDs or SLDs, and the way the demands are grouped and ungrouped (potentially several times) along the path. In a single-layer network, a simple strategy consisting in routing connections along the shortest path in terms of hops results in the lowest cost since the total number of required cross-connect ports is minimized. In a multi-layer network, the provisioning aimed at minimizing cost is less obvious since demands can be deviated from their shortest paths in order to create long, efficiently utilized high-rate connections to reduce the overall network cost. The provisioning process becomes even more complex when the demands are SxDs, instead of permanent connections, since the time-disjointness among demands needs to be taken into account in order to reuse network resources as much as possible.

16.3.1 Electrical Grooming

The characteristic of the SEDs lies in the fact that their flow is smaller than the capacity of a lightpath. This particularity is taken into account by use of electrical aggregation. The grooming of multiple SEDs consists in multiplexing their electrical flows onto the same Grooming Lightpath (GL) in an EXC. Grooming requires all the considered SEDs to share at least one common fiber link and all of them to be active during a common period of time. In order to explain the rationale of electrical traffic grooming, let us consider the simple case of the aggregation of two SEDs into a GL. These two SEDs, referred to as δ_1 and δ_2 , are characterized by the tuples $(s_1, d_1, n_1, \alpha_1, \beta_1)$ and $(s_2, d_2, n_2, \alpha_2, \beta_2)$, respectively.

Figure 16.2 illustrates the network topology that supports δ_1 and δ_2 . Below the illustration of the network we have plotted the time diagram of δ_1 and δ_2 . Both requests are active from α_2 to β_1 and they have a common path between node g_1 and node g_2 .

When the network does not have grooming capabilities, each SED is routed using a separate GL between its source and its destination nodes, as illustrated in Fig. 16.3. In other words, δ_1 is routed using the GL φ_1 between s_1 and d_1 , while δ_2 is routed using the GL φ_2 between s_2 and d_2 . As a result, two optical channels are needed on each fiber link between nodes g_1 and g_2 .

On the other hand, when the network has grooming capabilities, δ_1 and δ_2 can be groomed together between nodes g_1 and g_2 if $n_1 + n_2 \leq C_\omega$ as illustrated in Fig. 16.4 (C_ω is the nominal capacity of a lightpath). This grouping creates an aggregated demand $\delta_a = (g_1, g_2, n_1 + n_2, \alpha_2, \beta_1)$ borne by GL φ_a . However, since δ_1 is also active during period $[\alpha_1, \alpha_2]$ and δ_2 is also active during period $[\beta_1, \beta_2]$, two additional requests must be considered: $\delta_f = (g_1, g_2, n_1, \alpha_1, \alpha_2)$ to take into account the remaining active period of δ_1 and $\delta_g = (g_1, g_2, n_2, \beta_1, \beta_2)$

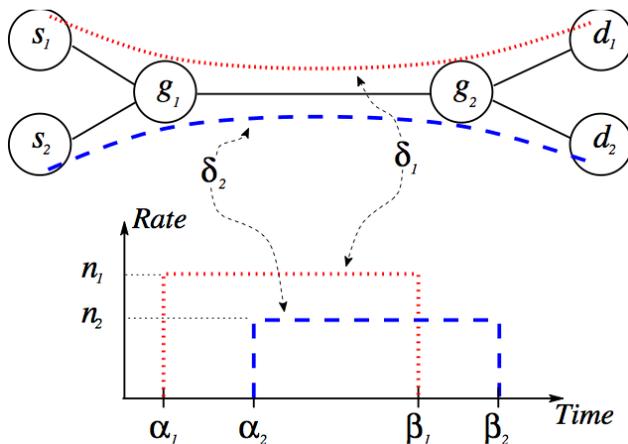


Fig. 16.2. Time diagram of two SEDs

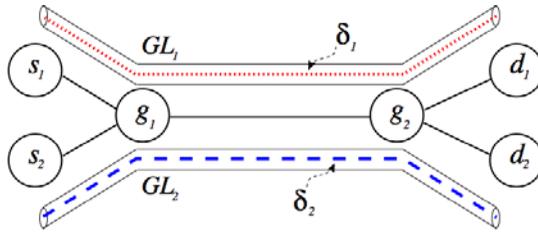


Fig. 16.3. Network without grooming capabilities

to take into account the remaining active period of δ_2 . In addition, we still need to take into account the non-common segments of the paths of δ_1 and δ_2 . This can be done by introducing four additional traffic requests δ_b , δ_c , δ_d , and δ_e :

- Request $\delta_b = (s_1, g_1, n_1, \alpha_1, \beta_1)$ carried by GL \wp_b represents a segment of the path of δ_1 totally disjoint from the path of δ_2 . It connects s_1 , the source node of δ_1 to the grooming node g_1 .
- Request $\delta_c = (g_2, d_1, n_1, \alpha_1, \beta_1)$ carried by GL \wp_c represents another segment of the path of δ_1 totally disjoint from the path of δ_2 . It connects the grooming node g_2 to d_1 , the destination node of δ_1 .

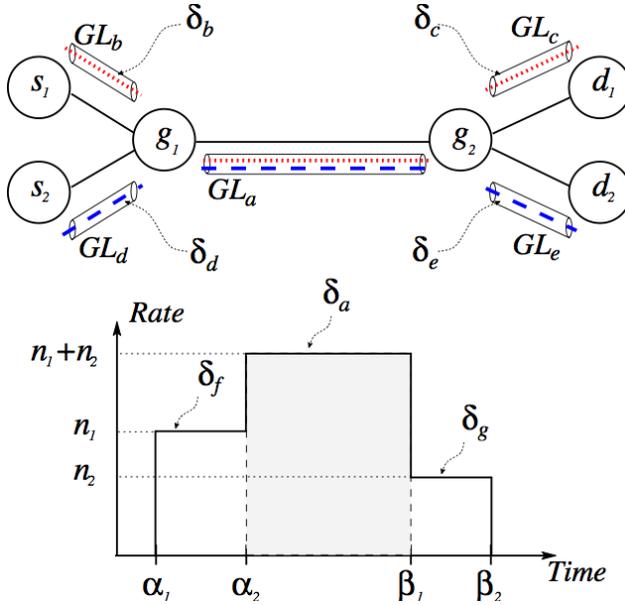


Fig. 16.4. Network with grooming capabilities

- Request $\delta_d = (s_2, g_1, n_2, \alpha_2, \beta_2)$ carried by GL φ_d represents a segment of the path of δ_2 totally disjoint from the path of δ_1 . It connects s_2 , the source node of δ_2 , to the grooming node g_1 .
- Request $\delta_e = (g_2, d_2, n_2, \alpha_2, \beta_2)$ represents another segment of the path of δ_2 totally disjoint from the path of δ_1 . It connects the grooming node g_2 to d_2 , the destination node of δ_2 .

To sum up, grooming together requests δ_1 and δ_2 creates the aggregated request δ_a and a set of six additional requests. Some of these requests ($\delta_b, \delta_c, \delta_d$, and δ_e) take into account the non-common segments of the paths of δ_1 and δ_2 . The remaining requests (δ_f and δ_g) take into account the non-common periods of time. Depending on the characteristics of the initial requests δ_1 and δ_2 , some of these additional requests may not exist. However, those which do exist form the set of marginal demands. As a result of the grooming process, when grooming two SEDs, the set of all the SEDs must be modified by adding the aggregated request δ_a and the set of marginal demands and by removing the original requests δ_1 and δ_2 .

At node g_1 , the two demands δ_1 and δ_2 must be switched from the WXC to the EXC in order to be groomed together. Hence two receiving r_3 electrical ports and two emitting o_3 ports are needed at this node. Once groomed together onto the the GL φ_a , the resulting grooming lightpath is switched back to the WXC using an emitting e_3 electrical port and a receiving o_3 optical port. Arriving at node g_2 , the GL φ_a is switched from the WXC to the EXC using an emitting o_3 optical port and a receiving r_3 electrical port. At the EXC of node g_2 , the initial requests δ_1 and δ_2 are rebuilt and each one is switched to its destination using two emitting e_3 electrical ports and two receiving o_3 optical ports. In brief, in order to groom these two requests, six additional electrical ports and six additional optical ports are needed. Figure 16.5 details the ports involved in the grouping and ungrouping operations at nodes g_1 and g_2 , respectively.

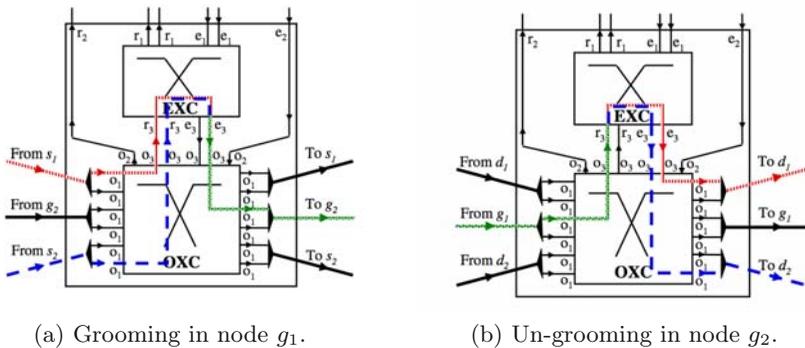


Fig. 16.5. Grooming operations in nodes g_1 and g_2

16.3.2 Optical Grooming

For the optical grooming problem, we consider a network of WXC/BXC nodes (no EXCs). We assume that WXCs provide full wavelength conversion. This simplifies wavelength assignment since the wavelength continuity constraint does not exist. As explained in Section 16.2, the lightpaths at the output (input) ports of the WXC are multiplexed (demultiplexed) into (from) wavebands which are directly added in (dropped from) a BXC. Thus, the WXCs are not directly connected to each other but through waveband-switching connections between BXCs. This means that lightpaths are instantiated over a logical topology of waveband-switching connections. These connections cannot be directly added in or dropped from the BXC. The number of WXC I/O ports is a multiple of the waveband size.

Figure 16.6 shows how three lightpaths can be instantiated over a logical topology formed by two waveband-switching connections, one between nodes 1 and 3 and another between nodes 3 and 4. Two lightpaths are added in the WXC of node 1 using the WXC add ports. They are then multiplexed into a waveband-switching connection, which is added into the BXC of the same node using one BXC add port. The resulting waveband-switching connection is multiplexed into a fiber. In node 2, the connection is demultiplexed from the incoming fiber, switched to an output port and multiplexed into a fiber. In node 3, the connection is demultiplexed from the incoming fiber and dropped using one BXC drop port. This connection is demultiplexed in order to retrieve the two lightpaths which are then switched from the input to the output ports of the node. The lightpaths are multiplexed together with a third lightpath added in node 3 into a waveband-switching connection which is added in the BXC of the node. The connection is multiplexed into a fiber. In node 4, the connection is demultiplexed from the incoming fiber and dropped. The three

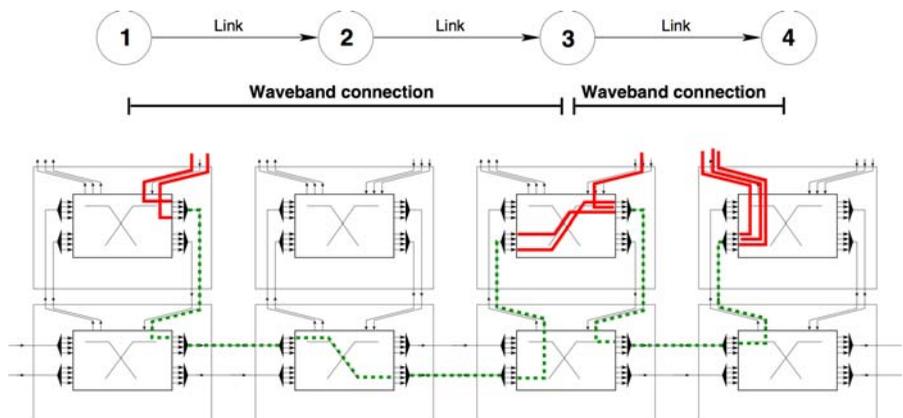


Fig. 16.6. A possible configuration of MG-OXCs used to set up three lightpaths

lightpaths are demultiplexed from the waveband-switching connection and dropped using the WXC drop ports of the node.

Waveband-switching introduces an intermediate network layer between the physical network and the SLDs. A logical topology of waveband-switching connections must be defined and mapped on the physical network. An SLD is instantiated by defining a path on the logical topology and assigning waveband-switching connections on each arc of the path to the SLD's lightpaths. In this context, grooming refers to the aggregation (and disaggregation) of lightpaths into waveband-switching connections of the logical topology. Thus, the optical grooming problem under consideration consists in defining a logical topology of waveband-switching connections, routing these connections on a physical network, routing a set of SLDs on the logical topology, and assigning the waveband-switching connections of the logical topology to the SLDs such that the network cost is minimal. The cost of the network is equal to the sum of the costs of the MG-OXC required to implement the solution. The problem is schematically illustrated by Fig. 16.7. An instance of the problem is defined by a set of SLDs and the topology of a physical network. The MG-OXC cost function models the cost of a MG-OXC as a function of the number of WXC and BXC ports of the MG-OXC (the function is defined in Equation 16.16). The network functional model defines the architecture of the switches, including the definition of the size of the wavebands (the number of lightpaths that can be groomed into a waveband-switching connection). The objective function defines the optimality criterion to be satisfied, for example, the minimization of the network cost. The solution to the problem consists of a set of Routed Scheduled Band Groups (RSBGs) and of an assignment of SLDs to RSBGs. An RSBG is similar to an SLD in that it is defined by a tuple $(s, d, n, \alpha, \beta, P)$ where s and d are the source and destination BXC of the RSBG, n is the number of waveband-switching connections in the group, α and β are the set-up and tear-down dates of the RSBG, and P is the route in the physical network. The set-up date α (tear-down date β) of an RSBG is the earliest set-up (latest tear-down) date of any of the SLDs assigned to this RSBG.

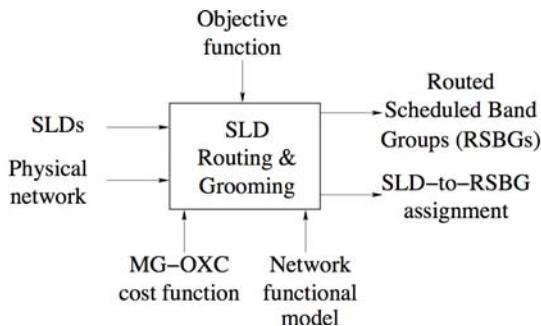


Fig. 16.7. Schematic representation of the optical grooming problem

16.4 Mathematical Formulation

In this section, we develop the mathematical formulations used to quantify the network cost in the electrical and optical grooming problems. The notations in Sections 16.4.1 and 16.4.2 are independent, that is, the symbols used in one subsection are not related to the symbols of the other.

16.4.1 Electrical Grooming

The design of a grooming-capable network is formally stated below. The inputs to the problem are:

- A physical topology represented by an arc-weighted symmetrical directed graph $G = (V; E; w)$. $V = \{v_1, v_2, \dots, v_N\}$ is the set of network nodes and $E = \{e_1, e_2, \dots, e_F\}$ is the set of physical links interconnecting the nodes. Nodes correspond to EXC/WXC multi-layer switches and links correspond to the fibers of the network. Though links are unidirectional, we assume that there are equal number of fibers linking two nodes in different directions. Links are assigned weights $w : E \mapsto \mathbb{R}^+$ which may correspond to the physical length or cost of the links. $N = |V|$ and $F = |E|$ represent the number of nodes and links in the network, respectively.
- A set $\mathcal{D} = \{\delta_1, \delta_2, \dots, \delta_M\}$ of M SxD requests to be routed on the network. These requests can be divided into a set $\mathcal{D}^L = \{\delta_1^L, \delta_2^L, \dots, \delta_{M^L}^L\}$ of M^L SLD requests and a set $\mathcal{D}^E = \{\delta_1^E, \delta_2^E, \dots, \delta_{M^E}^E\}$ of M^E SED requests. SLDs (δ_i^L) and SEDs (δ_i^E) are represented by a tuple $(s_i, d_i, n_i, \alpha_i, \beta_i)$ where $s_i, d_i \in V$ are the source and destination nodes of the demand, n_i is the traffic rate requested by the connection, and α_i and β_i are its setup and tear-down dates. For SLDs, n_i is an integer, while for SEDs n_i is a real number smaller than 1.
- A set of available routes $\mathcal{P}_i = \{P_{i,1}, P_{i,2}, \dots, P_{i,K}\}$ connecting the source node s_i to the destination node d_i of request δ_i^L or δ_i^E . For each request, we compute beforehand K alternate shortest paths in terms of effective length connecting its source node to its destination node using the algorithm described in [14] (if as many paths exist, otherwise we consider the available ones).

The set-up and tear-down dates of all these requests are to be taken into account because they represent time instants when a change in the traffic flow is observed. Let \mathcal{E} be the ordered set grouping the set-up and tear-down dates of all the requests in the set \mathcal{D} . Because some requests may have the same set-up and tear-down dates, the number $\mathbb{E} = |\mathcal{E}|$ of time instants in \mathcal{E} is less than or equal to $2 \times M$.

$$\begin{aligned} \mathcal{E} &= \left(\bigcup_{i=1}^M \alpha_i \right) \cup \left(\bigcup_{i=1}^M \beta_i \right) \\ &= \{\epsilon_1, \epsilon_2, \dots, \epsilon_{\mathbb{E}}\} \text{ such as } \epsilon_1 < \epsilon_2 < \dots < \epsilon_{\mathbb{E}} \end{aligned} \quad (16.1)$$

In the case of SEDs, a physical path $P_{i,\rho_i} \in \mathcal{P}_i$ is assigned in a first step to the request δ_i^E ($1 \leq i \leq M^E$). P_{i,ρ_i} is the ρ_i^{th} shortest path between s_i and d_i ($1 \leq \rho_i \leq K$). Once the physical path P_{i,ρ_i} is chosen, we determine in a second step the intermediate nodes of P_{i,ρ_i} in which δ_i^E must be switched to the electrical level for grooming. In this way, we obtain the set of grooming lightpaths that the request traverses in the logical topology.

A physical path P_{i,ρ_i} , composed of ℓ_{i,ρ_i} hops, has $(\ell_{i,\rho_i} - 1)$ intermediate nodes and can be decomposed into $\mathbb{L}_{i,\rho_i} = 2^{\ell_{i,\rho_i} - 1}$ different sets of grooming lightpaths. For example, the physical path $P_{i,\rho_i} = 1 - 2 - 3 - 4$ between nodes 1 and 4 can be decomposed into the following $2^{3-1} = 4$ different sets of grooming lightpaths:

- $L_{i,\rho_i,1} = \{[1 - 2 - 3 - 4]\}$
- $L_{i,\rho_i,2} = \{[1 - 2], [2 - 3 - 4]\}$
- $L_{i,\rho_i,3} = \{[1 - 2 - 3], [3 - 4]\}$
- $L_{i,\rho_i,4} = \{[1 - 2], [2 - 3], [3 - 4]\}$

$L_{i,\rho_i,\lambda_{i,\rho_i}}$ is the $\lambda_{i,\rho_i}^{\text{th}}$ possible decomposition of the physical path P_{i,ρ_i} into a set of grooming lightpaths ($1 \leq \lambda_{i,\rho_i} \leq \mathbb{L}_{i,\rho_i}$). Let $\mathcal{L}_{i,\rho_i} = \{L_{i,\rho_i,1}, L_{i,\rho_i,2}, \dots, L_{i,\rho_i,\mathbb{L}_{i,\rho_i}}\}$ be the set of all possible decompositions of the path P_{i,ρ_i} into its inherent sets of grooming lightpaths.

The solution of the optimization problem consists in assigning to the request δ_i^E a physical path $P_{i,\rho_i} \in \mathcal{P}_i$ and a set of grooming lightpaths $L_{i,\rho_i,\lambda_{i,\rho_i}} \in \mathcal{L}_{i,\rho_i}$. This solution is represented by $\Pi_{\mathcal{D}^E,\rho,\lambda} = \{(P_{1,\rho_1}, L_{1,\rho_1,\lambda_{1,\rho_1}}), (P_{2,\rho_2}, L_{2,\rho_2,\lambda_{2,\rho_2}}), \dots, (P_{M^E,\rho_{M^E}}, L_{M^E,\rho_{M^E},\lambda_{M^E,\rho_{M^E}}})\}$.

Our objective is to satisfy all the requests at the lowest cost. The network cost is expressed in terms of the number of electrical ports and of optical ports used and can be obtained by means of mathematical operations based on matrix algebra.

Computation of Network Cost

For an instance, $\Pi_{\mathcal{D}^E,\rho,\lambda} = \{(P_{1,\rho_1}, L_{1,\rho_1,\lambda_{1,\rho_1}}), (P_{2,\rho_2}, L_{2,\rho_2,\lambda_{2,\rho_2}}), \dots, (P_{M^E,\rho_{M^E}}, L_{M^E,\rho_{M^E},\lambda_{M^E,\rho_{M^E}}})\}$ of the network design problem under SED traffic requests, we define the following matrices:

- **Request Matrix:** The request matrix, denoted by $\underline{\Delta}^E = \{\delta_{i,t}^E\}$, represents the SED traffic requests over time. $\underline{\Delta}^E$ is a $[M^E \times \mathbb{E}]$ matrix. An element $\delta_{i,t}^E$ of such a matrix is a binary value ($\delta_{i,t}^E \in \{0, 1\}$) specifying the presence ($\delta_{i,t}^E = 1$) or absence ($\delta_{i,t}^E = 0$) of the SED request $\delta_i^E(s_i, d_i, \alpha_i, \beta_i, n_i)$ at time instant ϵ_t .

$$\delta_{i,t}^E = \begin{cases} 1 & \text{if } \alpha_i \leq \epsilon_t < \beta_i, \\ 0 & \text{otherwise.} \end{cases} \tag{16.2}$$

- **Source/Destination Matrix:** The source matrix, denoted by $\underline{S}^E = \{s_{i,n}^E\}$, represents the source nodes of the SED requests. The destination matrix, denoted by $\underline{D}^E = \{d_{i,n}^E\}$, represents the destination nodes of the SED requests. \underline{S}^E and \underline{D}^E are $[M^E \times N]$ matrices. An element $s_{i,n}^E$ ($d_{i,n}^E$) of such a matrix is a binary value specifying if the SED request $\delta_i^E(s_i, d_i, \alpha_i, \beta_i, n_i)$ has node v_n as source (destination) node or not.

$$s_{i,n}^E = \begin{cases} 1 & \text{if } s_i = v_n, \\ 0 & \text{otherwise.} \end{cases} \quad d_{i,n}^E = \begin{cases} 1 & \text{if } d_i = v_n, \\ 0 & \text{otherwise.} \end{cases} \quad (16.3)$$

As shown in Section 16.3.1, when grooming two SED requests $\delta_{i_1}^E$ and $\delta_{i_2}^E$, the set \mathcal{D}^E of all the SEDs must be modified by adding the aggregated request and the set of marginal demands, and by removing the original requests $\delta_{i_1}^E$ and $\delta_{i_2}^E$. Consequently, the number of SED requests as well as their characteristics are modified. Let $\tilde{\mathcal{D}}^E = \{\tilde{\delta}_1^E, \tilde{\delta}_2^E, \dots, \tilde{\delta}_{M^E}^E\}$ be the set of the \tilde{M}^E groomed SED requests obtained at the end of the grooming optimization process. Each groomed SED request $\tilde{\delta}_i^E \in \tilde{\mathcal{D}}^E$ is routed over a unique path \tilde{P}_i^E . As each element $\tilde{\delta}_i^E$ can be either a single low-speed connection δ_j^E or the concatenation of two or several low-speed connections $\tilde{\delta}_i^E = \uplus_j \delta_j^E$, its corresponding path is determined according to the grooming lightpath sets of its lower-speed connections ($\tilde{P}_i^E \in L_{j,\rho_j,\lambda_j,\rho_j}$).

Based on the new SED set $\tilde{\mathcal{D}}^E$, we define the following additional matrices:

- **Groomed Request Matrix:** The groomed request matrix, denoted by $\underline{\tilde{\Delta}}^E = \{\tilde{\delta}_{i,t}^E\}$, represents the groomed SED traffic requests over time. $\underline{\tilde{\Delta}}^E$ is a $[\tilde{M}^E \times \mathbb{E}]$ matrix. An element $\tilde{\delta}_{i,t}^E$ of such a matrix is an integer number due to the fact that the groomed SED demand $\tilde{\delta}_i^E(\tilde{s}_i, \tilde{d}_i, \tilde{\alpha}_i, \tilde{\beta}_i, \tilde{n}_i)$ can be the concatenation of several requests of smaller rates. Consequently, it could require more than one lightpath to carry its traffic load. $\tilde{\delta}_{i,t}^E$ is equal to the smallest integer greater than or equal to the traffic required by the groomed SED request $\tilde{\delta}_i^E$ and specifies the number of lightpaths used by this request at time instant ϵ_t .

$$\tilde{\delta}_{i,t}^E = \begin{cases} \lceil \tilde{n}_i \rceil & \text{if } \tilde{\alpha}_i \leq \epsilon_t < \tilde{\beta}_i, \\ 0 & \text{otherwise.} \end{cases} \quad (16.4)$$

- **Groomed Routing Matrix:** The groomed routing matrix, denoted by $\underline{\tilde{R}}^E = \{\tilde{r}_{i,f}^E\}$, represents the use of the physical links by the groomed SED requests. $\underline{\tilde{R}}^E$ is a $[\tilde{M}^E \times F]$ matrix. An element $\tilde{r}_{i,f}^E$ of such a matrix is a binary value ($\tilde{r}_{i,f}^E \in \{0, 1\}$) specifying if the physical path \tilde{P}_i^E assigned to the groomed SED request $\tilde{\delta}_i^E(\tilde{s}_i, \tilde{d}_i, \tilde{\alpha}_i, \tilde{\beta}_i, \tilde{n}_i)$ passes through the physical link e_f ($\tilde{r}_{i,f}^E = 1$) or not ($\tilde{r}_{i,f}^E = 0$).

$$\tilde{r}_{i,f}^E = \begin{cases} 1 & \text{if } e_f \in \tilde{P}_i^E, \\ 0 & \text{otherwise.} \end{cases} \quad (16.5)$$

- **Groomed Source/Destination Matrix:** The groomed source matrix, denoted by $\tilde{\underline{S}}^E = \{\tilde{s}_{i,n}^E\}$, represents the source nodes of the groomed SED requests. The groomed destination matrix, denoted by $\tilde{\underline{D}}^E = \{\tilde{d}_{i,n}^E\}$, represents the destination nodes of the groomed SED requests. $\tilde{\underline{S}}^E$ and $\tilde{\underline{D}}^E$ are $[\tilde{M}^E \times N]$ matrices. An element $\tilde{s}_{i,n}^E$ ($\tilde{d}_{i,n}^E$) of such a matrix is a binary value specifying if the groomed SED request $\tilde{\delta}_i^E(\tilde{s}_i, \tilde{d}_i, \tilde{\alpha}_i, \tilde{\beta}_i, \tilde{\eta}_i)$ has node v_n as source (destination) node or not.

$$\tilde{s}_{i,n}^E = \begin{cases} 1 & \text{if } \tilde{s}_i = v_n, \\ 0 & \text{otherwise.} \end{cases} \quad \tilde{d}_{i,n}^E = \begin{cases} 1 & \text{if } \tilde{d}_i = v_n, \\ 0 & \text{otherwise.} \end{cases} \quad (16.6)$$

By multiplying the transposed request matrix $\underline{\Delta}^{E^T}$ by the source matrix \underline{S}^E a new matrix $\underline{\Psi}^{E,Em} = \{\psi_{t,n}^{E,Em}\}$ is obtained. $\underline{\Psi}^{E,Em}$, of dimension $[\mathbb{E} \times N]$, represents the use of sub-wavelength ‘add’ ports at the nodes. An element $\psi_{t,n}^{E,Em}$ of such a matrix is an integer value representing the number of emitting e_1 electrical ports in use at node v_n at each instant ϵ_t . For a given node v_n , let $v_n^{E,Em}$ be the maximum value of the nodes’ add ports $\psi_{t,n}^{E,Em}$ evaluated over the whole observation period. $v_n^{E,Em}$ gives the number of emitting e_1 electrical ports installed at the node v_n . The number $\gamma_{e_1}^E$ of emitting e_1 electrical ports in the network is the sum of $v_n^{E,Em}$ over all the network nodes.

$$\underline{\Psi}^{E,Em} = \underline{\Delta}^{E^T} \times \underline{S}^E \quad (16.7a)$$

$$v_n^{E,Em} = \max_{1 \leq t \leq \mathbb{E}} \psi_{t,n}^{E,Em} \quad (16.7b)$$

$$\gamma_{e_1}^E = \sum_{n=1}^N v_n^{E,Em} \quad (16.7c)$$

Similarly, the number $\gamma_{r_1}^E$ of receiving r_1 electrical ports in the network is computed as follows:

$$\underline{\Psi}^{E,Re} = \underline{\Delta}^{E^T} \times \underline{D}^E \quad (16.8a)$$

$$v_n^{E,Re} = \max_{1 \leq t \leq \mathbb{E}} \psi_{t,n}^{E,Re} \quad (16.8b)$$

$$\gamma_{r_1}^E = \sum_{n=1}^N v_n^{E,Re} \quad (16.8c)$$

By multiplying the transposed groomed request matrix $\tilde{\Delta}^{E^T}$ by the groomed routing matrix \tilde{R}^E , a new matrix $\tilde{\Phi}^E = \{\tilde{\phi}_{t,f}^E\}$ is obtained. $\tilde{\Phi}^E$, of dimension $[\mathbb{E} \times F]$, represents the traffic load over time carried by the links. An element $\tilde{\phi}_{t,f}^E$ of such a matrix is an integer value representing the number of WDM optical channels carried by link e_f at each instant e_t . For a given link e_f , let $\tilde{\varphi}_f^E$ be the maximum value of this traffic $\tilde{\phi}_{t,f}^E$ evaluated over the whole observation period. $\tilde{\varphi}_f^E$ gives the number of WDM optical channels used on this link. As the o_1 optical ports go by pair, one port at each end of a WDM channel, $\tilde{\varphi}_f^E$ represents also the number of o_1 optical port pairs installed at both ends of the link e_f . The number $\gamma_{o_1}^E$ of o_1 optical ports in the network is twice the sum of $\tilde{\varphi}_f^E$ over all the network links. The maximum value of $\tilde{\varphi}_f^E$ ($\forall f \setminus 1 \leq f \leq F$) represents the congestion in the network, i.e., the number of active channels on the most loaded link.

$$\tilde{\Phi}^E = \tilde{\Delta}^{E^T} \times \tilde{R}^E \tag{16.9a}$$

$$\tilde{\varphi}_f^E = \max_{1 \leq t \leq \mathbb{E}} \tilde{\phi}_{t,f}^E \tag{16.9b}$$

$$\gamma_{o_1}^E = 2 \times \sum_{f=1}^F \tilde{\varphi}_f^E \tag{16.9c}$$

By multiplying the transposed groomed request matrix $\tilde{\Delta}^{E^T}$ by the groomed source matrix \tilde{S}^E a new matrix $\tilde{\Psi}^{E,Em} = \{\tilde{\psi}_{t,n}^{E,Em}\}$ is obtained. $\tilde{\Psi}^{E,Em}$, of dimension $[\mathbb{E} \times N]$, represents the use of switching ports from the EXC to the WXC at the nodes. An element $\tilde{\psi}_{t,n}^{E,Em}$ of such a matrix is an integer value representing the number of emitting e_3 electrical ports in use at node v_n at each instant e_t . For a given node v_n , let $\tilde{v}_n^{E,Em}$ be the maximum value of the nodes' switching ports $\tilde{\psi}_{t,n}^{E,Em}$ evaluated over the whole observation period. $\tilde{v}_n^{E,Em}$ gives the number of emitting e_3 electrical ports installed at the node v_n . The number $\gamma_{e_3}^E$ of emitting e_3 electrical ports in the network is the sum of $\tilde{v}_n^{E,Em}$ over all the network nodes.

$$\tilde{\Psi}^{E,Em} = \tilde{\Delta}^{E^T} \times \tilde{S}^E \tag{16.10a}$$

$$\tilde{v}_n^{E,Em} = \max_{1 \leq t \leq \mathbb{E}} \tilde{\psi}_{t,n}^{E,Em} \tag{16.10b}$$

$$\gamma_{e_3}^E = \sum_{n=1}^N \tilde{v}_n^{E,Em} \tag{16.10c}$$

Similarly, the number $\gamma_{r_3}^E$ of receiving r_3 electrical ports in the network is computed as follows:

$$\underline{\tilde{\Psi}}^{E,Re} = \underline{\tilde{\Delta}}^{E^T} \times \underline{\tilde{D}}^E \quad (16.11a)$$

$$\tilde{v}_n^{E,Re} = \max_{1 \leq t \leq \mathbb{E}} \tilde{\psi}_{t,n}^{E,Re} \quad (16.11b)$$

$$\gamma_{r_3}^E = \sum_{n=1}^N \tilde{v}_n^{E,Re} \quad (16.11c)$$

Consequently, the number $\gamma_{o_3}^E$ of o_3 optical ports in the network is equal to:

$$\gamma_{o_3}^E = \gamma_{e_3}^E + \gamma_{r_3}^E \quad (16.12)$$

As a set of SEDs does not have any SLD component to be added or dropped at a node, the number of e_2/r_2 electrical ports and the number of o_2 optical ports are null.

$$\gamma_{e_2}^E = \gamma_{r_2}^E = \gamma_{o_2}^E = 0 \quad (16.13)$$

Finally, the total number ϱ^E of optical ports and the total number ε^E of electrical ports are given by:

$$\begin{aligned} \varrho^E &= \gamma_{o_1}^E + \gamma_{o_2}^E + \gamma_{o_3}^E \\ &= \gamma_{o_1}^E + \gamma_{o_3}^E \end{aligned} \quad (16.14a)$$

$$\begin{aligned} \varepsilon^E &= \gamma_{e_1}^E + \gamma_{e_2}^E + \gamma_{e_3}^E + \gamma_{r_1}^E + \gamma_{r_2}^E + \gamma_{r_3}^E \\ &= \gamma_{e_1}^E + \gamma_{e_3}^E + \gamma_{r_1}^E + \gamma_{r_3}^E \end{aligned} \quad (16.14b)$$

κ being the ratio of the cost of an electrical port to the cost of an optical port, the cost ζ^E of routing the SED requests is then expressed as:

$$\zeta^E = \varrho^E + \kappa \times \varepsilon^E \quad (16.15)$$

16.4.2 Optical Grooming

$G = (V, E, w)$

is an arc-weighted symmetrical directed graph with vertex set $V = \{v_1, v_2, \dots, v_N\}$, arc set $E = \{e_1, e_2, \dots, e_F\}$, and arc weight function $w : E \rightarrow \mathbb{R}_+$. The graph represents a physical telecommunications network. The set V corresponds to the nodes of the network and the set of arcs E to the links interconnecting the nodes. Function w corresponds to the physical length or to the cost of the links (defined, for example, by the network operator).

$N = |V|$, $F = |E|$, K , L , B

are, respectively, the number of nodes and links in the network, the maximum number of possible alternate paths for each demand, the maximum number of possible layouts for each path and the size of a waveband-switching connection

(expressed in number of lightpaths). All the waveband-switching connections in the network have the same size.

$$P = (x_0, x_1, \dots, x_z)$$

describes a *path* in G that is composed of z links $(x_0, x_1), (x_1, x_2), \dots, (x_{z-1}, x_z)$ where the $(x_{i-1}, x_i) \in E$ are all distinct (i.e., the path is loop free).

$$\Delta = \{\delta_1, \delta_2, \dots, \delta_M\}$$

is a set of M Scheduled Lightpath Demands (SLDs), where:

$$\delta_i = (s_i, d_i, n_i, \alpha_i, \beta_i)$$

is a tuple representing SLD number i ; $s_i, d_i \in V$ are the source and destination nodes of the demand, n_i is the number of requested lightpaths, and α_i and β_i are the set-up and tear-down dates.

$$(G, \Delta)$$

is a pair representing an instance of the optical grooming problem.

$$P_{i,k} = (x_0^{i,k}, x_1^{i,k}, \dots, x_{z_i,k}^{i,k}), 1 \leq i \leq M, 1 \leq k \leq K$$

is the k -th alternate path for SLD δ_i from $x_0^{i,k} = s_i$ to $x_{z_i,k}^{i,k} = d_i$. For the purposes of this model, we compute the K physically⁴ shortest paths (if so many exist) for each demand using the algorithm defined in [14]. However, the paths might be defined according to any other criterion (i.e., the function w may map any other value than the length of the links).

$$C : \Pi_\Delta \rightarrow \mathbb{R}_+$$

is the function that computes the cost of an admissible solution $\pi_{\rho,\nu,\Delta}$ to the optical grooming problem (explained below). In order to formalize this function we define the following additional notations.

$$\theta = (\theta_{ij})$$

is a $\{0, 1\}^{M \times M}$ upper triangular matrix; θ_{ij} , $i \leq j$, indicates whether the SLDs δ_i and δ_j overlap in time ($\theta_{ij} = 1$) or not ($\theta_{ij} = 0$). By definition $\theta_{ii} = 1$, $1 \leq i \leq M$, and $\theta_{ij} = 0$ for $i > j$. This matrix expresses the temporal interdependence between the SLDs.

$$\tau = (\tau_{ij}) = \text{diag}(n_i)$$

is a diagonal matrix in which $\tau_{ii} = n_i$, $1 \leq i \leq M$, that is, τ_{ii} is the number of lightpaths required by the SLD δ_i .

$$\mathcal{I} = (\mathcal{I}_{ij})$$

is a $\{0, 1\}^{N \times F}$ matrix; \mathcal{I}_{ij} indicates whether vertex v_i is the termination vertex of arc e_j ($\mathcal{I}_{ij} = 1$) or not ($\mathcal{I}_{ij} = 0$).

⁴ The function w maps the length of each link.

$$\mathcal{O} = (\mathcal{O}_{ij})$$

is a $\{0, 1\}^{N \times F}$ matrix; \mathcal{O}_{ij} indicates whether vertex v_i is the source vertex of arc e_j ($\mathcal{O}_{ij} = 1$) or not ($\mathcal{O}_{ij} = 0$).

$$\mathcal{T} = (t_{ij})$$

is a $\{0, 1\}^{M \times N}$ matrix; t_{ij} indicates whether vertex v_j is the source node of SLD δ_i ($t_{ij} = 1$) or not ($t_{ij} = 0$).

$$\mathcal{U} = (u_{ij})$$

is a $\{0, 1\}^{M \times N}$ matrix; u_{ij} indicates whether vertex v_j is the destination node of SLD δ_i ($u_{ij} = 1$) or not ($u_{ij} = 0$).

$$\mathcal{D} = \theta \cdot \tau \cdot \mathcal{T} = (\mathcal{D}_{ij})_{1 \leq i \leq M, 1 \leq j \leq N},$$

$$\mathcal{D}_j^* = \max_{1 \leq i \leq M} \{\mathcal{D}_{ij}\}$$

\mathcal{D}^* is an N -dimensional vector; \mathcal{D}_j^* is the maximum number of simultaneously active lightpaths (i.e., time-overlapping lightpaths) originating at node v_i .

$$\mathcal{E} = \theta \cdot \tau \cdot \mathcal{U} = (\mathcal{E}_{ij})_{1 \leq i \leq M, 1 \leq j \leq N},$$

$$\mathcal{E}_j^* = \max_{1 \leq i \leq M} \{\mathcal{E}_{ij}\}$$

\mathcal{E}^* is an N -dimensional vector; \mathcal{E}_j^* is the maximum number of simultaneously active lightpaths (i.e., time-overlapping lightpaths) terminating at node v_i .

$$\psi : \mathbb{N} \rightarrow \mathbb{R}_+$$

is the function that determines the cost of a cross-connect (either a WXC or a BXC) as a function of its number of ports. In this work, we consider the function:

$$\psi(x) = a + bx^c \quad a, b \in \mathbb{R}_+, c \in [1, 2[. \quad (16.16)$$

The function captures various technology-specific factors that have an effect on the cost of a cross-connect. The parameter a represents the fixed cost of installing the switch (shelf, power and ventilation systems, etc.). The parameter b represents the cost of a port. Finally, the parameter c accounts for the impact on the cost of the increasing implementation complexity of switching matrices with a large number of ports. This parameter is limited to take values smaller than 2 because, for higher values, it is in general more economical to stack multiple small switches than building a large one when a significant number of ports is required.

$$\lambda = \{\lambda^n\}$$

is a partition of the set of arcs describing path P . We call λ a *layout* and an element λ^n of this partition a *subpath*. For example, a layout of path $(x_0, x_1, x_2, x_3, x_4)$ is $\lambda = \{(x_0, x_1, x_2), (x_2, x_3, x_4)\}$. This layout has subpaths $\lambda^1 = (x_0, x_1, x_2)$ and $\lambda^2 = (x_2, x_3, x_4)$. Another layout is $\{(x_0, x_1), (x_1, x_2, x_3, x_4)\}$. The elements of a subpath must be contiguous arcs of P .

$$\Lambda_{i,k} = \{\lambda_{i,k,j}\}, 1 \leq i \leq M, 1 \leq k \leq K, 1 \leq j \leq L$$

is the set of layouts available for path $P_{i,k}$; $\lambda_{i,k,j}$ is the j -th layout of the path. We assume that there are at most L different layouts defined for a path.

$$\pi_{\rho,\nu,\Delta} = ((P_{1,\rho_1}, \lambda_{1,\rho_1,\nu_1}), (P_{2,\rho_2}, \lambda_{2,\rho_2,\nu_2}), \dots, (P_{M,\rho_M}, \lambda_{M,\rho_M,\nu_M})),$$

$$\rho \in \{1, \dots, K\}^M, \nu \in \{1, \dots, L\}^M$$

is called an *admissible optical grooming solution* for Δ . ρ and ν are M -dimensional vectors whose elements can take values in $[1, K]$ and $[1, L]$, respectively. $\pi_{\rho,\nu,\Delta}$ is fully characterized by the pair (ρ, ν) . An admissible optical grooming solution $\pi_{\rho,\nu,\Delta}$ defines for each SLD, a path $P_{i,k}$ and a layout $\lambda_{i,k,j}$ of this path. A subpath λ^n can be part of several layouts defined in $\pi_{\rho,\nu,\Delta}$. We call the association of λ^n to the subset $\Delta' \subseteq \Delta$ of SLDs whose layouts in $\pi_{\rho,\nu,\Delta}$ share λ^n , a Routed Scheduled Band Group (RSBG). Thus, besides a path and a layout for each SLD, an admissible optical grooming solution $\pi_{\rho,\nu,\Delta}$ also defines a set of RSBGs.

$$\Pi_{\Delta} = \{\pi_{\rho,\nu,\Delta}, \rho \in \{1, \dots, K\}^M, \nu \in \{1, \dots, L\}^M\}$$

is the set of all admissible optical grooming solutions for Δ . Its cardinality is $|\Pi_{\Delta}| = (KL)^M$ (assuming that K paths and L layouts are available for each path; otherwise, $|\Pi_{\Delta}| < (KL)^M$). The set represents the solution space of an optical grooming problem instance (G, Δ) .

$$\mathcal{S}_{\rho,\nu,\Delta} = \bigcup_{i=1}^M \lambda_{i,\rho_i,\nu_i}$$

is the set of all subpaths used in the admissible optical grooming solution $\pi_{\rho,\nu,\Delta}$. The cardinality of the set is denoted $S = |\mathcal{S}_{\rho,\nu,\Delta}|$. For the sake of simplicity, we note \mathcal{S} instead of $\mathcal{S}_{\rho,\nu,\Delta}$.

$$\mathcal{A} = (a_{ij})$$

is a $\{0, 1\}^{M \times S}$ matrix; a_{ij} indicates whether SLD δ_i uses subpath \mathcal{S}_j ($a_{ij} = 1$) or not ($a_{ij} = 0$) in admissible optical grooming solution $\pi_{\rho,\nu,\Delta}$.

$$\mathcal{B} = (b_{ij})$$

is a $\{0, 1\}^{F \times S}$ arc-subpath incidence matrix; b_{ij} indicates whether arc e_i is part of subpath \mathcal{S}_j ($b_{ij} = 1$) or not ($b_{ij} = 0$).

$$\eta = \theta \cdot \tau \cdot \mathcal{A} \cdot \mathcal{B}^T = (\eta_{ij})$$

is a $\mathbb{N}^{M \times F}$ matrix; η_{ij} indicates the number of time-overlapping lightpaths on link e_j between SLD δ_i and SLDs $\delta_{i+1}, \delta_{i+2}, \dots, \delta_M$.

$$\eta^* = (\eta_j^*)_{1 \leq j \leq F}, \eta_j^* = \left\lceil \frac{1}{B} \max_{1 \leq i \leq M} \eta_{ij} \right\rceil$$

is an F -dimensional vector; η_j^* indicates the number of wavebands required on arc e_j .

$$\mathcal{IN} = \mathcal{I} \cdot \eta^*$$

is an N -dimensional vector; \mathcal{IN}_i indicates the number of input ports required in the BXC of node v_i to implement admissible optical grooming solution $\pi_{\rho, \nu, \Delta}$.

$$\mathcal{OUT} = \mathcal{O} \cdot \eta^*$$

is an N -dimensional vector; \mathcal{OUT}_i indicates the number of output ports required in the BXC of node v_i to implement admissible optical grooming solution $\pi_{\rho, \nu, \Delta}$.

$$\mathcal{G} = (g_{ij})$$

is a $\{0, 1\}^{N \times S}$ matrix; g_{ij} indicates whether vertex v_i is the source of subpath \mathcal{S}_j ($g_{ij} = 1$) or not ($g_{ij} = 0$).

$$\mathcal{H} = (h_{ij})$$

is a $\{0, 1\}^{N \times S}$ matrix; h_{ij} indicates whether vertex v_i is the termination of subpath \mathcal{S}_j ($h_{ij} = 1$) or not ($h_{ij} = 0$).

$$\mathcal{J} = \theta \cdot \tau \cdot \mathcal{A} \cdot \mathcal{G}^T = (\mathcal{J}_{ij})_{1 \leq i \leq M, 1 \leq j \leq N},$$

$$\mathcal{J}_j^* = \left\lceil \frac{1}{B} \max_{1 \leq i \leq M} \mathcal{J}_{ij} \right\rceil$$

\mathcal{J}^* is an N -dimensional vector; \mathcal{J}_j^* is the number of waveband connections added at the BXC of node v_j .

$$\mathcal{K} = \theta \cdot \tau \cdot \mathcal{A} \cdot \mathcal{H}^T = (\mathcal{K}_{ij})_{1 \leq i \leq M, 1 \leq j \leq N},$$

$$\mathcal{K}_j^* = \left\lceil \frac{1}{B} \max_{1 \leq i \leq M} \mathcal{K}_{ij} \right\rceil$$

\mathcal{K}^* is an N -dimensional vector; \mathcal{K}_j^* is the number of waveband connections dropped at the BXC of node v_j .

$$\mathcal{Q} \in \mathbb{R}_+^N,$$

$$\mathcal{Q}_i = \psi_b(\max(\mathcal{IN}_i, \mathcal{OUT}_i) + \max(\mathcal{J}_i^*, \mathcal{K}_i^*))$$

is an N -dimensional vector; \mathcal{Q}_i indicates the cost of the BXC required at node v_i to implement admissible optical grooming solution $\pi_{\rho, \nu, \Delta}$.

$$\mathcal{R} \in \mathbb{R}_+^N,$$

$$\mathcal{R}_i = \psi_w(\max(\mathcal{D}_i^*, \mathcal{E}_i^*) + B \max(\mathcal{J}_i^*, \mathcal{K}_i^*))$$

is an N -dimensional vector; \mathcal{R}_i indicates the cost of the WXC required at node v_i to implement admissible optical grooming solution $\pi_{\rho, \nu, \Delta}$. Note that the number of input/output ports, $B \max(\mathcal{J}_i^*, \mathcal{K}_i^*)$, is a multiple of the waveband size B . The parameters defining the function ψ_w may be different from those defining the function ψ_b .

The cost function $C : \Pi_{\Delta} \rightarrow \mathbb{R}_+$ is defined as:

$$C(\pi_{\rho, \nu, \Delta}) = \sum_{i=1}^N (Q_i + R_i). \tag{16.17}$$

16.5 Illustrative Examples

16.5.1 Electrical Grooming

In this subsection, we present a small example in order to highlight the reduction of network cost provided by the electrical grooming functionality. We considered the three 3 SEDs of Table 16.1. Their time diagram is shown in Fig. 16.8. To emphasize the benefit of grooming we use a fixed routing solution in which each SED is routed along its shortest path. Emitting e_1 electrical ports and receiving r_1 electrical ports are only needed to add and drop the traffic requests. Consequently, the number of such ports is independent of the routing and grooming solution. When the network does not have grooming capabilities (Fig. 16.9(a)), the SEDs δ_1 and δ_2 are routed over distinct lightpaths. Thus, we need two WDM optical channels on each of the 3–4 and the 4–7 fiber links. As a result, this routing and grooming solution requires a total of 9 WDM optical channels. Let us recall that the number of required o_1 optical ports is twice the number of WDM optical channels used. On the other hand, when the network has grooming capabilities (Fig. 16.9(b)), the SEDs δ_1 and δ_2 can be groomed together since both requests are active from 11:00 to 13:00 and they share a common path between nodes 3 and 7. Therefore, a grooming lightpath is created between these nodes using only one WDM optical channel on each of the 3–4 and the 4–7 fiber links. This grooming

Table 16.1. Characteristics of 3 SEDs

δ_i	s_i	d_i	α_i	γ_i	n_i
δ_1	2	8	8:00	14:40	$0.25 \times C_{\omega}$
δ_2	3	7	11:00	13:00	$0.25 \times C_{\omega}$
δ_3	2	6	17:00	19:30	$0.25 \times C_{\omega}$

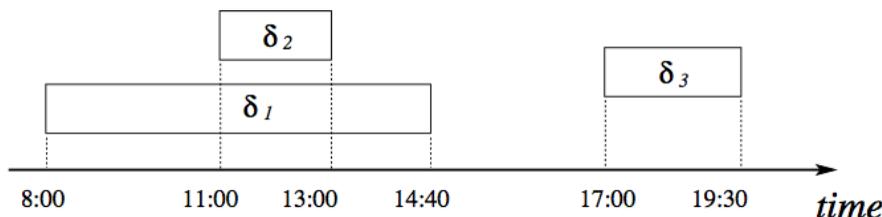


Fig. 16.8. Associated time diagram of the 3 SEDs

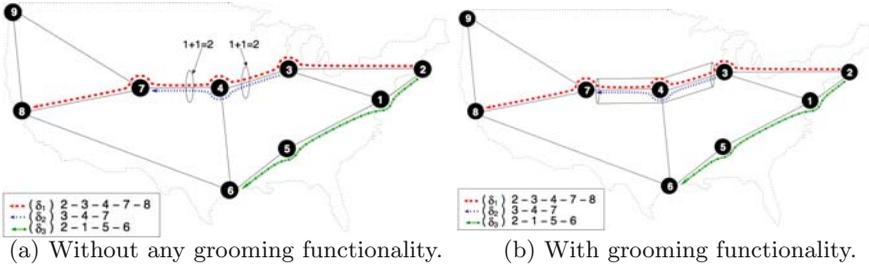


Fig. 16.9. Provisioning in a network of the 3 SEDs of Table 16.1

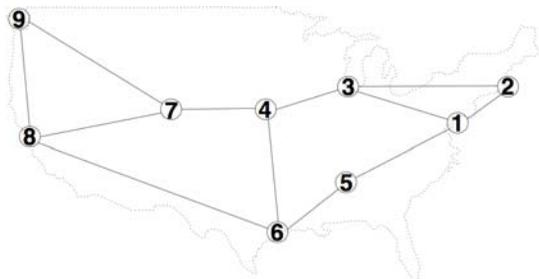
lightpath is used to carry simultaneously the two SEDs δ_1 and δ_2 between nodes 3 and 7. In our example, the SED δ_1 still needs to be routed using two additional lightpaths; the first one between nodes 2 and 3, and the second one between nodes 7 and 8. As a result, only 7 WDM optical channels are needed for this new routing and grooming solution. However, it should be noted that, in order to aggregate SEDs δ_1 and δ_2 at node 3, one additional emitting o_3 optical port and one additional receiving r_3 electrical port are needed on this node. Similarly, to de-aggregate these two SEDs at node 7, one additional emitting e_3 electrical port and one additional receiving o_3 optical port are needed. Depending on the relative costs of the different o_1 , o_3 , e_3 , and r_3 ports, the new routing and grooming solution can be more or less economical than the first solution.

16.5.2 Optical Grooming

In this subsection, we present an example to illustrate how the mathematical model of Section 16.4.2 is used to represent an instance of the optical grooming problem and to quantify the cost of a solution to this problem.

We study the problem instance of Fig. 16.10 which defines a set Δ of $M = 2$ SLDs and a graph G representing a physical network. Additionally, we choose the paths and layouts defined in Table 16.2 ($K = 2$ and $L = 3$) and define a band size of $B = 4$. For the sake of simplicity, we choose functions $\psi_w(x) = x$ and $\psi_b(x) = x$ to represent the cost of switches, i.e., we choose $a = 0$, $b = 1$ and $c = 1$ in Equation (16.16). Figure 16.11 illustrates the admissible optical grooming solution $\pi_{\rho,\nu,\Delta}$ with vectors $\rho = (1, 1)$ and $\nu = (3, 3)$ for the instance of Fig. 16.10. Path $P_{1,1}$ and layout $\lambda_{1,1,3}$ are selected for SLD δ_1 and path $P_{2,1}$ and layout $\lambda_{2,1,3}$ are selected for SLD δ_2 . The figure also shows the set of Routed Scheduled Band Groups (RSBGs) defined by the solution.

The following matrices are used to compute the cost of the chosen admissible optical grooming solution:



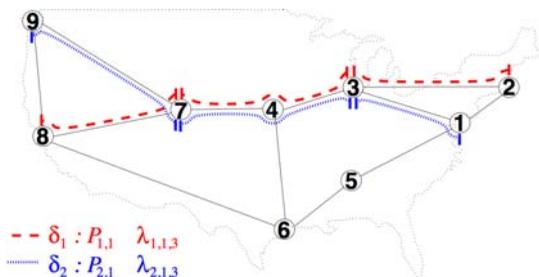
SLD	s	d	n	α	β
δ_1	2	8	8	08:00	14:00
δ_2	1	9	10	11:00	13:00

(a) Graph G representing a physical network. (b) Set Δ of $M = 2$ SLDs.

Fig. 16.10. An instance (G, Δ) of the optical grooming problem

Table 16.2. Chosen paths and layouts for the instance of Fig. 16.10

SLD	Path ($P_{i,k}$)	Layout ($\lambda_{i,k,j}$)
δ_1	$P_{1,1} = (2, 3, 4, 7, 8)$	$\lambda_{1,1,1} = \{(2, 3, 4, 7, 8)\}$
		$\lambda_{1,1,2} = \{(2, 3, 4) (4, 7, 8)\}$
		$\lambda_{1,1,3} = \{(2, 3) (3, 4, 7) (7, 8)\}$
	$P_{1,2} = (2, 1, 5, 6, 8)$	$\lambda_{1,2,1} = \{(2, 1, 5, 6, 8)\}$
		$\lambda_{1,2,2} = \{(2, 1, 5) (5, 6, 8)\}$
		$\lambda_{1,2,3} = \{(2, 1) (1, 5, 6) (6, 8)\}$
δ_2	$P_{2,1} = (1, 3, 4, 7, 9)$	$\lambda_{2,1,1} = \{(1, 3, 4, 7, 9)\}$
		$\lambda_{2,1,2} = \{(1, 3, 4) (4, 7, 9)\}$
		$\lambda_{2,1,3} = \{(1, 3) (3, 4, 7) (7, 9)\}$
	$P_{2,2} = (1, 5, 6, 8, 9)$	$\lambda_{2,2,1} = \{(1, 5, 6, 8, 9)\}$
		$\lambda_{2,2,2} = \{(1, 5, 6) (6, 8, 9)\}$
		$\lambda_{2,2,3} = \{(1, 5) (5, 6, 8) (8, 9)\}$



Subpath (λ^n)	$\Delta' \subseteq \Delta$
(1, 3)	$\{\delta_2\}$
(2, 3)	$\{\delta_1\}$
(3, 4, 7)	$\{\delta_1, \delta_2\}$
(7, 8)	$\{\delta_1\}$
(7, 9)	$\{\delta_2\}$

-- $\delta_1 : P_{1,1} \lambda_{1,1,3}$
 --- $\delta_2 : P_{2,1} \lambda_{2,1,3}$

(a) Paths and layouts.

(b) Routed Scheduled Band Groups (RSBGs).

Fig. 16.11. An admissible optical grooming solution $\pi_{\rho,\nu,\Delta}$ to the instance of Fig. 16.10

$$\theta = \begin{matrix} & \delta_1 & \delta_2 \\ \delta_1 & \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \end{matrix} \quad \gamma = \begin{matrix} & \delta_1 & \delta_2 \\ \delta_1 & \begin{pmatrix} 8 & 0 \\ 0 & 10 \end{pmatrix} \end{matrix}$$

$$\mathcal{S}_{\rho, \nu, \Delta} = \{(1, 3), (2, 3), (3, 4, 7), (7, 8), (7, 9)\}$$

$$\mathcal{A} = \begin{matrix} & s_1 & s_2 & s_3 & s_4 & s_5 \\ \delta_1 & \begin{pmatrix} 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \end{pmatrix} \end{matrix} \quad \mathcal{B} = \begin{matrix} & s_1 & s_2 & s_3 & s_4 & s_5 \\ e_{13} & \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix}$$

$$\eta = \theta \cdot \gamma \cdot \mathcal{A} \cdot \mathcal{B}^T = \begin{pmatrix} 10 & 8 & 18 & 18 & 8 & 10 \\ 10 & 0 & 10 & 10 & 0 & 10 \end{pmatrix} \quad \eta^* = \begin{pmatrix} 3 \\ 2 \\ 5 \\ 5 \\ 2 \\ 3 \end{pmatrix}$$

$$\mathcal{I} = \begin{matrix} & e_{13} & e_{23} & e_{34} & e_{47} & e_{78} & e_{79} \\ v_1 & \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix} \quad \mathcal{O} = \begin{matrix} & e_{13} & e_{23} & e_{34} & e_{47} & e_{78} & e_{79} \\ \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \end{matrix}$$

$$\mathcal{IN} = \mathcal{I} \cdot \eta^* = (0, 0, 5, 5, 0, 0, 5, 2, 3)$$

$$\mathcal{OUT} = \mathcal{O} \cdot \eta^* = (3, 2, 5, 5, 0, 0, 5, 0, 0)$$

$$\mathcal{G} = \begin{matrix} & s_1 & s_2 & s_3 & s_4 & s_5 \\ v_1 & \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \end{matrix} \quad \mathcal{H} = \begin{matrix} & s_1 & s_2 & s_3 & s_4 & s_5 \\ v_1 & \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix}$$

$$\begin{aligned}
 \mathcal{T} &= \begin{matrix} & v_1 & v_2 & v_3 & \dots & v_9 \\ \delta_1 & \left(\begin{array}{cccccc} 0 & 1 & 0 & \dots & 0 \\ 1 & 0 & 0 & \dots & 0 \end{array} \right) & \mathcal{U} &= \begin{matrix} & v_1 & \dots & v_7 & v_8 & v_9 \\ \delta_1 & \left(\begin{array}{cccccc} 0 & \dots & 0 & 1 & 0 \\ 0 & \dots & 0 & 0 & 1 \end{array} \right) \end{matrix} \\
 \mathcal{D} &= \theta \cdot \gamma \cdot \mathcal{T} = \begin{pmatrix} 10 & 8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 10 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} & \mathcal{D}^* &= (10 \ 8 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0) \\
 \mathcal{E} &= \theta \cdot \gamma \cdot \mathcal{U} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 8 & 10 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 10 \end{pmatrix} & \mathcal{E}^* &= (0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 8 \ 10) \\
 \mathcal{J} &= \theta \cdot \gamma \cdot \mathcal{A} \cdot \mathcal{G}^T = \begin{pmatrix} 10 & 8 & 18 & 0 & 0 & 0 & 18 & 0 & 0 \\ 10 & 0 & 10 & 0 & 0 & 0 & 10 & 0 & 0 \end{pmatrix} & \mathcal{J}^* &= (3 \ 2 \ 5 \ 0 \ 0 \ 0 \ 5 \ 0 \ 0) \\
 \mathcal{K} &= \theta \cdot \gamma \cdot \mathcal{A} \cdot \mathcal{H}^T = \begin{pmatrix} 0 & 0 & 18 & 0 & 0 & 0 & 18 & 8 & 10 \\ 0 & 0 & 10 & 0 & 0 & 0 & 10 & 0 & 10 \end{pmatrix} & \mathcal{K}^* &= (0 \ 0 \ 5 \ 0 \ 0 \ 0 \ 5 \ 2 \ 3) \\
 \mathcal{Q} &= (6 \ 4 \ 10 \ 5 \ 0 \ 0 \ 10 \ 4 \ 6) & \mathcal{R} &= (22 \ 16 \ 20 \ 0 \ 0 \ 0 \ 20 \ 16 \ 22)
 \end{aligned}$$

The cost of the admissible optical grooming solution $\pi_{\rho, \nu, \Delta}$ according to (16.17) is:

$$C(\pi_{\rho, \nu, \Delta}) = 28 + 20 + 30 + 5 + 0 + 0 + 30 + 20 + 28 = 161. \quad (16.18)$$

16.6 Conclusion

This chapter presented mathematical formulations to quantify the cost of provisioning scheduled connections in multi-layer optical networks. The specific cases of electrical grooming in EXC/WXC networks and optical grooming in WXC/BXC networks were considered. The formulations may serve as a basis to assess the costs and benefits of introducing an additional network layer in an existing network, or to determine the impact on the network-wide cost of a change in the technologies used to implement a cross-connect. The use of scheduled connections make it possible to take into account dynamic changes on the traffic load in the cost/benefit analysis of multi-layer networks.

The formulations could be extended to compare the impact on network cost of alternative resiliency approaches such as link or path protection, dedicated or shared protection, etc. Another extension could be the consideration of the uncertainty about the tear-down date of SEDs or SLDs, since connections terminating later than their expected tear-down date might be seizing network resources allocated to demands that were originally time-disjoint with respect to the seizing connection. It is reasonable to expect that, the longer a connection lasts in the network, the higher the uncertainty about its actual tear-down date.

References

1. J. Kuri, N. Puech, M. Gagnaire, E. Dotaro and R. Douville, "Routing and Wavelength Assignment of Scheduled Lightpath Demands," *IEEE Journal on Selected Areas in Communications*, Vol. 21, N. 8, pp. 1231–1240, October 2003.
2. J. Kuri, "Optimization Problems in WDM Optical Transport Networks with Scheduled Lightpath Demands," *Ph.D. Thesis*, Ecole Nationale Supérieure des Télécommunications. Paris, France. September 2003.
3. N. Skorin-Kapov, "Heuristic algorithms for the routing and wavelength assignment of scheduled lightpath demands in optical networks," *IEEE Journal on Selected Areas in Communications*, Vol. 24, NO. 8, August 2006.
4. H. G. Ahn, T.-J. Lee, M. Y. Chung and H. Choo, "RWA on Scheduled Lightpath Demands in WDM Optical Transport Networks with Time Disjoint Paths," *Proceedings of the International Conference on Information Networking (ICOIN 2005)*, Jeju, Korea, February 2005.
5. C. V. Saradhi and M. Gurusami, "Graph Theoretic Approaches for Routing and Wavelength Assignment of Scheduled Lightpath Demands in WDM Optical Networks," *Proceedings of the 2nd International Conference on Broadband Networks (Broadnets 2005)*, Boston, MA, October 2005.
6. J. Kuri, N. Puech, M. Gagnaire and E. Dotaro, "Routing and Wavelength Assignment of Scheduled Lightpath Demands in a WDM Optical Transport Network," *Proceedings of the First International Conference in Optical Communications and Networks (ICOON 2002)*, Singapur, November 2002.
7. T. Li and B. Wang, "On optimal survivability design in WDM optical networks under a scheduled traffic model," *Proceedings of the 5th International Workshop on Design of Reliable Communication Networks (DRCN 2005)*, Ischia, Italy, October 2005.
8. T. Li and B. Wang, "On Optimal Survivability Design under a Scheduled Traffic Model in Wavelength-Routed Optical Mesh Networks," *Proceedings of the 4th Annual Communication Networks and Services research Conference (CNSR 2006)*, Moncton, Canada, May 2006.
9. C. V. Saradhi, L. K. Wei and M. Gurusami, "Provisioning Fault-Tolerant Scheduled Lightpath Demands in WDM Mesh Networks," *Proceedings of the First International Conference on Broadband Networks (BROADNETS 2004)*, San Jose, CA, October 2004.
10. E. A. Doumith and M. Gagnaire, "Traffic Routing in Multi-Layer Optical Network Considering Rerouting and Grooming Strategies," *Proceedings of the IEEE Global Communications Conference (Globecom)*, San Francisco, CA, December 2006.
11. E. A. Doumith, M. Gagnaire, O. Audouin, and R. Douville, "Network Nodes Dimensioning Assuming Electrical Traffic Grooming in an Hybrid OXC/EXC WDM Network," *Proceedings of the 2nd IEEE/Create-Net International Conference on Broadband Networks (BroadNets'05)*, Boston, MA, 2005.
12. E. A. Doumith and M. Gagnaire, "Traffic Grooming in Multi-Layer WDM Networks: Meta-heuristics versus Sequential Algorithms," *Proceedings of the IEEE Eunice Open European Summer School*, Stuttgart, Germany, September 2006.

13. E. A. Doumith, M. Gagnaire, O. Audouin and R. Douville, "From Network Planning to Traffic Engineering for Optical VPN and Multi-Granular Random Demands," *Proceedings of the 2nd IEEE/Create-Net International Conference on Broadband Networks (BroadNets'05)*, Phoenix, AZ, April 2006.
14. D. Eppstein, "Finding the k Shortest Paths," *IEEE Symposium on Foundations of Computer Science*, pp. 154–165, 1994.

All-Optical Traffic Grooming

Volkan Kaman, Henrik N. Poulsen, Roger J. Helkey, John E. Bowers

17.1 Introduction

Next-generation optical communication networks are expected to exceed aggregate capacities of hundreds of terabits per second to support the recent growth in commercial broadband access fueled by applications such as Video-on-Demand (VoD) and Voice-over-IP (VoIP), and to meet the high-performance demands of grid-computing networks for research consortiums and military applications [15, 26, 31, 32]. The continuous increase in need for higher bandwidth has also necessitated in parallel a continuous reduction in cost per bit, which has paved the way for the all-optical revolution. As higher bandwidth with smaller footprint and lower power consumption scales better with lower cost optics, the past 30 years have seen electronics get gradually pushed out of the core communication network to the lower bandwidth edge user interfaces. The advantage of transitioning to all-optical communication is most evident first with the introduction of the erbium-doped fiber amplifier (EDFA) in the early 1990s. The EDFA not only increased the all-optical reach of bandwidth independent signals by eliminating costly optical-electrical-optical (OEO) repeaters, but it also revolutionized point-to-point communication system capacity by allowing widespread implementation of dense wavelength division multiplexing (DWDM). The second confirmation of the benefit of all-optical communication is the introduction of optical add-drop multiplexers (OADM) at metro and backbone network nodes in the early 2000s. Similarly to the EDFA, the OADM allowed the removal of costly in-line OEO transponders by providing all-optical bypass of express traffic. Furthermore, in realizing the concept of the all-optical core ring network, the OADM provides significant optics savings by driving the grooming electronics switch to the edge of the network for sub-wavelength processing of the local add-drop traffic.

While the EDFA and OADM have allowed for the realization of today's all-optical point-to-point systems and core ring networks, the current shift to higher data rates of 40 Gb/s SONET, 100 Gb Ethernet and beyond opens up the possibility for further cost-saving all-optical innovations in high-capacity

edge and core networks. This chapter reviews the next-generation all-optical traffic grooming technologies for both the core and edge networks. Section 17.2 reviews the state-of-the-art smart optical switching technologies for all-optical wavelength grooming in the core network as these networks evolve to interconnected rings and mesh architectures. Section 17.3 reviews all-optical circuit grooming, or optical time division multiplexing (OTDM), for aggregating low-speed sub-wavelength circuits into wavelengths between the core and edge networks. Finally, Section 17.4 reviews all-optical edge grooming techniques such as optical packet switching (OPS) and optical burst switching (OBS) for reducing costly electronics in edge grooming switches and IP routers, and Section 17.5 concludes the chapter.

17.2 All-Optical Wavelength Grooming

The aggregate bandwidth capacity increase in next-generation metro and long haul core optical networks requires a considerable increase in DWDM fiber link capacity, which is achieved by increasing both the number of wavelengths and the data rate per wavelength using spectrally efficient modulation formats as recently demonstrated with a 25.6 Tb/s transmission experiment [6]. These high-capacity DWDM links are interconnected into dynamically reconfigurable all-optical mesh networks with efficient wavelength provisioning and optimized use of network elements for both capital and operational cost savings. This agile photonic network is enabled by the next-generation of remotely reconfigurable OADM (ROADM) with express traffic switching and local add-drop wavelength grooming capabilities. The ROADM should support multiple degree nodes, which is defined as the number of DWDM fiber network ports terminating at the node. It should also allow each add-drop port to have fully flexible access to the network to ensure efficient network utility of each access transponder. The latter is of high significance in a multi-degree ROADM design as transponders and regenerators dominate the overall core network cost and should be minimized to the lowest possible required by the network [5, 31]. The key enablers for all-optical wavelength grooming of core network traffic are the dramatic savings in (i) cost per express wavelength switching in the optical domain by removing in-line OEO transponders and electronic switches, and (ii) cost per access wavelength switching by removing redundant edge OEO transponders through flexible optical add-drop.

Traditional degree-two nodes in ring networks have relied upon static OADMs that can transit and add-drop only fixed wavelengths without reconfigurability. These OADMs are either based on a three-port WDM filter approach (Fig. 17.1a) or a pair of concatenated optical demultiplexers (DMUX) and multiplexers (MUX) for splitting and recombining all DWDM wavelengths (Fig. 17.1b). Next-generation degree-two ROADMs are based on wavelength blockers (WB) using liquid crystals (LC), or one-dimensional micro-electro-mechanical-system (1-D MEMS) switches, or integrated planar

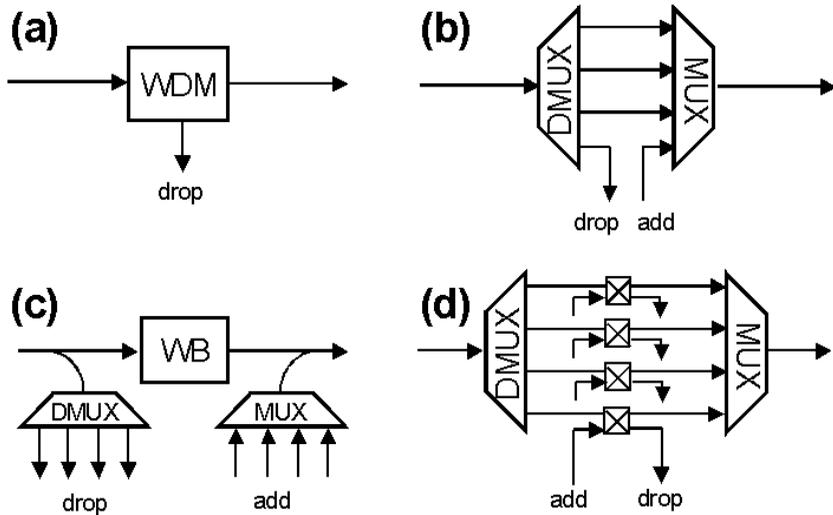


Fig. 17.1. Schematics of degree-two static OADM's based on (a) a WDM filter, (b) concatenated DMUX and MUX; and ROADM's based on (c) WB's, and (d) concatenated DMUX and MUX with 2×2 switches for add-drop access

lightwave circuits (PLC), where all techniques process and power balance express traffic at a per wavelength basis [18, 23]. Add-drop access for both 1-D MEMS and LC-based ROADMs is achieved by tapping the aggregate DWDM signal before and after the WB and separating the channels using fixed wavelength D/MUXs (Fig. 17.1c). The PLC-based ROADM has the advantage of utilizing the already integrated express D/MUXs for fixed wavelength add-drop access by further integrating per channel 2×2 switches onto the PLC device (Fig. 17.1d). While degree-two ROADMs provide reconfigurable all-optical wavelength grooming, fully flexible add-drop access requires the use of additional tunable filters and optical switches, the cost of which can be prohibitive for high add-drop ratios.

As next-generation networks evolve into interconnected rings and mesh architectures, multi-degree ROADM nodes with flexible add-drop capability gain high significance. The add-drop flexibility criteria that eliminate additionally redundant and expensive electronics with cheap optics are summarized as follows [10]:

1. *directionless access* with any add-drop port to any DWDM network port connectivity,
2. *colorless access* with any wavelength-transparent add-drop port to any DWDM network port connectivity, and
3. *contentionless access* between all same-wavelength add-drop ports.

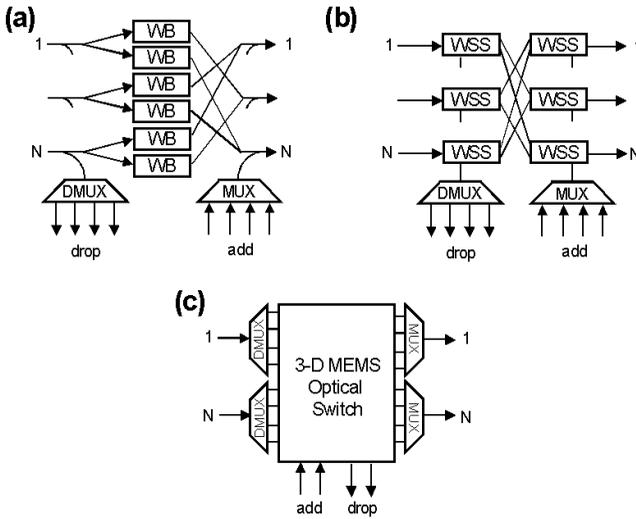


Fig. 17.2. Schematics of multi-degree ROADM architectures (a) broadcast-and-select using WBs, (b) optical cross-connect using WSSs, (c) WSXC using 3-D MEMS optical switch

Multi-degree ROADMs can be realized using three distinct technologies as summarized in Fig. 17.2: (a) a broadcast-and-select architecture using WBs, (b) an optical cross-connect (or broadcast-and-select) architecture using wavelength selective switches (WSS), and (c) a wavelength-selective cross-connect (WSXC) architecture using a large-scale 3-D MEMS optical switch.

The broadcast-and-select architecture shown in Fig. 17.2a is based on the same 1-D MEMS or LC WB technologies as discussed for degree-two ROADMs [30]. While passive optical splitters are used to broadcast and combine each incoming and outgoing DWDM network port, the WBs select the desired wavelengths to be transmitted. An inherent disadvantage of this architecture is that as the degree of the node N increases, the required number of WBs increases quadratically to $N(N - 1)$ while the ROADM express through loss increases significantly as well. This multi-degree ROADM also physically separates the express wavelength switching from the add-drop switching by tapping the DWDM signals before and after the core express switch. While the degree based modularity of this ROADM is attractive from a network restoration perspective, this also severely complicates the ability to provide flexible add-drop access with the aforementioned flexibility criteria.

A more attractive alternative to WB's is to use 1-D MEMS or LC based $1 : K$ WSS devices (typically with $K < 10$), which distribute an incoming DWDM signal into any of the desired K outputs [2, 20]. A multi-degree ROADM node can be achieved by cascading two WSS devices per network port and inter-connecting $N - 1$ WSS ports from one network port to the others

as shown in Fig. 17.2b [1]. While any of the remaining $K - N + 1$ WSS ports can be used for colorless local add-drop access of a single transponder, the limited number of available WSS ports requires optical aggregation (using D/MUX, splitters, and EDFAs) of the transponders before accessing the WSS. Furthermore, directionless and wavelength contentionless add-drop access requires further optical processing at the edge, which becomes undesirable for high-degree nodes with high add-drop ratios [13]. While the preceding described the optical cross-connect architecture using $2N$ WSS devices with a fixed through loss, typical ROADM implementations use the broadcast-and-select architecture by cascading an input splitter with a single WSS device per network port [1]. This architecture has the advantage that the number of required WSS devices per node is reduced to N ; however, the challenges with flexible add-drop access remain with an additional express insertion loss drawback.

A multi-degree ROADM can also be implemented with a WSXC architecture using a large-scale 3-D MEMS optical switch combined with a set of D/MUXs (Fig. 17.2c) [14]. While non-blocking 3-D MEMS switches with low optical insertion losses are commercially available with 320 ports, these matrix switches have also been shown to scale beyond 1,000 ports [9, 17]. Wavelength switching between DWDM network ports is achieved by directing each wavelength separately to a MEMS port through the use of modular or integrated D/MUX at the DWDM network ports. Furthermore, fully agile add-drop access is realized due to the colorless and non-blocking nature of each MEMS port. A transponder or regenerator can connect to the entire core network at any wavelength using a single MEMS port. Since the add-drop ports use the already existing express filtering devices, no additional elements are required for the add-drops, reducing cost compared to the preceding WSS architectures.

The three multi-degree ROADM technologies are compared in Fig. 17.3 in terms of WSS functionality and evolution. The first generation WSS is the WB with a single input and output. The limited number of ports on this 1:1 WSS not only results in a quadratic need of these devices in implementing a multi-degree ROADM, but also the absence of local access ports requires additional components with limited add-drop flexibility. The next-generation device is the 1: K WSS, which has a single input port and several output ports. The number of devices required to implement a multi-degree ROADM scales linearly with the number of network ports. While a limited number of the K ports can be used for colorless local access, full add-drop flexibility and high add-drop ratios requires additional components and complicated access architectures. Finally, the integrated WSXC realizes the multi-degree ROADM through a $N:N$ WSS device with additional ports available for fully flexible add and drop access at wavelength granularity with simple operational management. While the latter approach provides the most flexible multi-degree ROADM due to the non-blocking switching nature of each wavelength port, the former two architectures provide a modular growth possibility with less add-drop flexibility. While modular growth of 3-D MEMS switches has been proposed, it can be concluded that network port modularity and

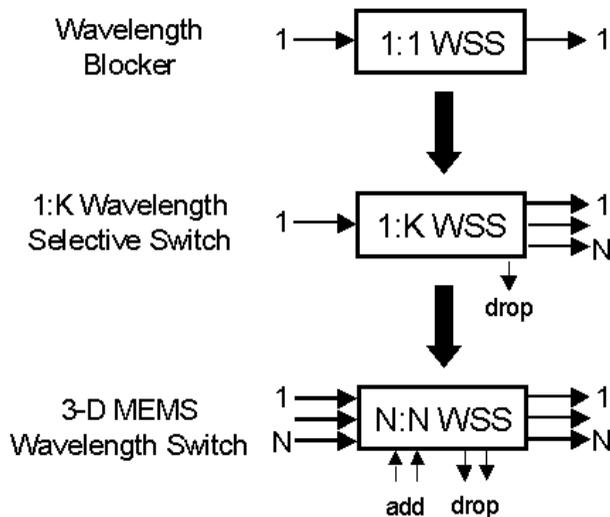


Fig. 17.3. Comparison of multi-degree ROADMs technologies in terms of WSS functionality and evolution

local add-drop flexibility of a multi-degree ROADMs are competing design criteria [13].

17.3 All-Optical Circuit Grooming

While DWDM is widely utilized to upgrade fiber bandwidth capacity, the cost, provisioning, and management of a huge number of wavelength channels can be highly problematic when considering dynamic mesh networks as opposed to point-to-point systems. Increasing the bit rate of each wavelength channel using traditional serial TDM circuit grooming while keeping the number of DWDM wavelengths to a cost effective and reasonably manageable number is therefore required to satisfy the bandwidth requirements of next-generation high-capacity networks. While TDM based on electronic grooming is currently being deployed at 40 Gb/s in commercial networks and demonstrated at 100 Gb/s in lab trials [19, 29], the speed and cost limitations of electronics will most likely require all-optical circuit grooming or OTDM techniques for achieving higher data rates of 160 Gb/s and beyond per wavelength [33]. As sub-wavelength multiplexing and demultiplexing are performed in the optical domain, OTDM systems have the advantage of needing opto-electronic component bandwidths only at the lower speed tributary rate. For example, a 160 Gb/s data rate can be achieved with commercially available low-cost 10 GHz devices [24]. Optical grooming or aggregation capability between the edge and core networks also allows for efficient utilization of resources for dynamic

sub-wavelength traffic. Furthermore, as lower speed circuits within a given wavelength may have different destinations, the ability to drop the desired circuits and insert new circuits in the available slots, or add-drop multiplexing, becomes a desirable functionality [27]. While traditional TDM requires electronic processing of all sub-wavelength traffic within a given wavelength, optical add-drop multiplexing using OTDM techniques can provide complete optical transparency to express sub-wavelength traffic from source to destination.

The basic principle of OTDM is illustrated in a point-to-point system in Fig. 17.4a. A typical OTDM transmitter consists of an optical pulse source that emits short pulses at the desired wavelength with a repetition rate of T , which is determined by the bit rate B at which electronics can follow. A passive $1:M$ splitter broadcasts these pulses into separate paths, where each pulse stream is independently data modulated and time delayed. All M pulse streams are then bit interleaved into a single serial data stream with an aggregate data rate of $B \times M$. The number of serial channels that can be packed into a single wavelength depends on the width of the optical pulses. On the OTDM receiver side, a $1:M$ splitter broadcasts the aggregate data stream to each of the M tributary detectors. A phase-locked optical gating device (DMUX) is utilized prior to the detector for extracting the desired sub-wavelength channel. The concept of add-drop multiplexing using the aforementioned OTDM techniques is shown in Fig. 17.4b. An optical gate extracts the desired channel to be dropped at the node while the rest of the channels are transmitted transparently. A new channel is then inserted in the available time slot.

While various OTDM technologies have been demonstrated in lab trials based on all-optical fiber techniques for ultra-high-speed single wavelength data rates [25, 33], the high number of devices required in OTDM sub-systems

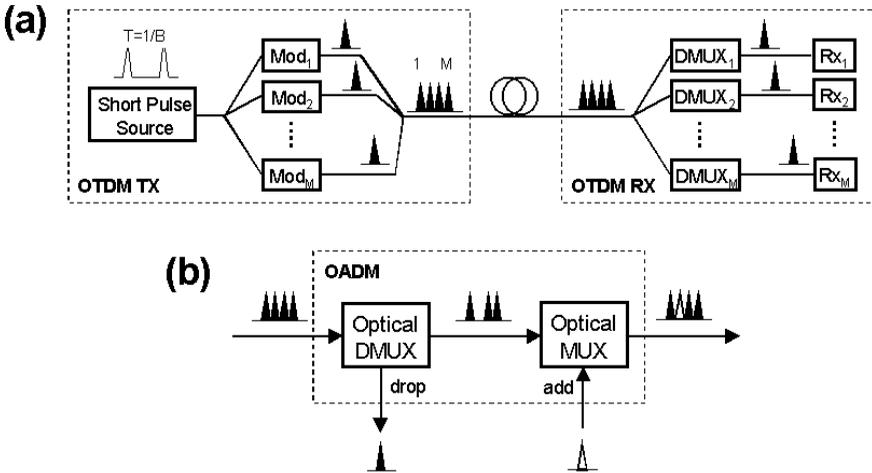


Fig. 17.4. (a) Basic schematic of an OTDM transmitter and receiver in a point-to-point system, (b) concept of OTDM optical add-drop multiplexing

makes photonic integration essential for commercial viability. Compact optical short pulse sources using a tandem of integrated electro-absorption modulators [8] and further integration with semiconductor optical amplifiers (SOA) for compensating losses have been demonstrated [22]. Further chip level integration of splitters, arrays of modulators and amplifiers on PLC platforms also show promise for commercial viability [16]. An OTDM demultiplexing receiver with an integrated electro-absorption modulator for optical gating followed by a detector has been demonstrated at 40 Gb/s [7]. Further integration of optical amplification has also been achieved for better receiver sensitivity performance [21]. Monolithically integrated SOA devices have been used for add-drop applications [28] while a single electro-absorption modulator has also been shown to simultaneously demultiplex and detect a single channel while transparently transmitting the express channels [12]. Using this capability, an alternative OTDM receiver architecture that imitates high-speed electrical receiver sub-systems with the benefit of operating at the tributary base rate has also been proposed [11]. The compact OTDM receiver eliminates the $1:M$ splitter and requires a single input fiber into a series of integrated and cascaded modulators to achieve high-speed serial-optical to direct tributary-speed parallel-electrical conversion. Apart from photonic integration, there are several other challenges that need to be considered for OTDM to have commercial penetration. The generation, multiplexing, and gating of such short pulses imposes very strict and challenging polarization, crosstalk, extinction ratio and timing with active signal processing requirements. Finally, further challenges of optical transmission of ultra-high data rates in dynamic networks will require thoughtful consideration.

17.4 All-Optical Edge Grooming

Present day networks are based on optical circuit switching with dedicated traffic paths carrying traffic from source to destination. As network bandwidth demands increase, DWDM systems with higher count channels and faster data rates are utilized to provide the required bandwidth capacity. While DWDM is a very effective means of using the bandwidth of the installed fiber base, such switching paradigms exhibit poor performance when sub-wavelength traffic is considered. As described in the preceding section, all-optical circuit grooming using OTDM techniques can help in aggregating several optical channels onto one wavelength. However, the limited granularity and static nature of OTDM switching does not render this as an efficient solution in edge networks. Therefore, traffic grooming at the edge of a network, which presently makes use of the fine granularity electronics can offer, is becoming increasingly important to improve the usage of the available optical resources and to increase the overall network utilization. In SONET and SDH networks, this is mainly accomplished by statistical multiplexing, where data from a large number of logical channels are carried on a single wavelength in the

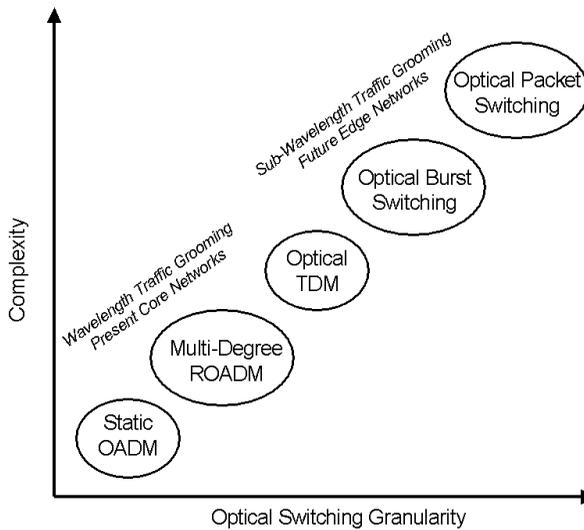


Fig. 17.5. Complexity of the various all-optical traffic grooming technologies

physical medium. The available bandwidth is dynamically allocated only towards active channels, which enables more devices to be connected than other multiplexing techniques. However, as services such as VoD and VoIP are being introduced, the requirements for low latency and high granularity are increasing. Due to the high cost and power consumption of high bit-rate electronics and with progressively more emphasis on transparency, all-optical edge grooming using OPS and OBS are becoming more and more interesting [34].

From a traffic utilization point of view, the highest gains are currently envisaged by using OPS since it can provide both the granularity and latency required for multi-media services [3, 35]. Instead of encapsulating each IP packet into larger SONET frames that are transmitted in a synchronous bit stream, with OPS each IP packet is transmitted by itself with an additional small header. Thus, IP packets arrive at the end destination asynchronously as they are not electronically processed and synchronized at each node in the network. While OPS provides highly efficient and granular switching, it also requires new technology at the edge that also permeates into the whole core network. The current lack of switches with nanosecond switching times and optical buffers are prohibitive factors for OPS networks with substantial development time required.

A hybrid between optical circuit and packet switching is OBS [4]. In contrast to OPS where packets are sent out one by one, OBS networks aggregate packets with the same end destination into larger packets, which are denoted as bursts. With significantly larger packets, the speed requirements of the switching elements in the core network are a lot less strict at the expense of decreased network utilization. The minimum length of a burst is determined

by the switching technology used and has an upper limit for the time allowed in which a burst can be assembled. As a burst cannot arbitrarily wait long enough for IP packets to arrive in assembling the burst, many bursts are likely to be timed out and transmitted without the minimum number of IP packets, in which case the network utilization is decreased. While OBS alleviates the issue of fast switching, it still shares the issue of optical buffers as in OPS.

The complexity of the various technologies discussed in this chapter for all-optical traffic grooming is summarized in Fig. 17.5 with respect to switching granularity.

17.5 Conclusion

In the quest for high-capacity and low-cost optical networks, traditional electronic traffic grooming is being replaced by more efficient all-optical technologies. Multi-degree ROADMs with wavelength grooming capabilities are already being deployed in commercial core networks while active research in OTDM, OPS, and OBS is underway for all-optical sub-wavelength grooming of edge traffic.

References

1. E.B. Basch, R. Egorov, S. Gringeri, and S. Elby. Architectural tradeoffs for reconfigurable dense wavelength-division multiplexing systems. *IEEE Journal in Selected Topics on Quantum Electronics*, 12:615–626, 2006.
2. G. Baxter, S. Frisken, D. Abakoumov, Z. Hao, I. Clarke, A. Bartos, and S. Poole. Highly programmable wavelength selective switch based on liquid crystal on silicon switching elements. In *Optical Fiber Communications Conference (OFC/NFOEC), paper OTuF2*, 2006.
3. D.J. Blumenthal, J.E. Bowers, L. Rao, H.-F. Chou, S. Rangarajan, and H.N. Poulsen. Optical signal processing for optical packet switching networks. *IEEE Communications Magazine*, 41:S23–S29, 2003.
4. Y. Chen, C. Qiao, and X. Yu. Optical burst switching: A new area in optical networking research. *IEEE Network Magazine*, 18(1):16–23, 2004.
5. O. Gerstel and H. Raza. Predeployment of resources in agile photonic networks. *IEEE Journal of Lightwave Technology*, 22:2236–2244, 2004.
6. A.H. Gnauck, G. Charlet, P. Tran, P. Winzer, C. Doerr, J. Centanni, E. Burrows, T. Kawanishi, T. Sakamoto, and K. Higuma. 25.6-Tb/s C+L-band transmission of polarization-multiplexed RZ-DQPSK signals. In *Optical Fiber Communications Conference (OFC/NFOEC), paper PDP19*, 2007.
7. V. Kaman, Y.J. Chiu, T. Liljeberg, S.Z. Zhang, and J.E. Bowers. A compact 40 Gbit/s demultiplexing receiver based on integrated tandem electroabsorption modulators. *IEEE Electronics Letters*, 36:1943–1944, 2000.
8. V. Kaman, Y.J. Chiu, T. Liljeberg, S.Z. Zhang, and J.E. Bowers. Integrated tandem traveling-wave electroabsorption modulators for > 100 Gbit/s OTDM applications. *IEEE Photonics Technology Letters*, 12:1471–1473, 2000.

9. V. Kaman, R.J. Helkey, and J.E. Bowers. Compact and scalable 3-D MEMS optical switches. *Journal of Optical Networks*, 6:19–24, 2007.
10. V. Kaman, R.J. Helkey, and J.E. Bowers. Multi-degree ROADMs with agile add-drop access. In *Photonics in Switching Conference, paper TuA2*, 2007.
11. V. Kaman, A.J. Keating, S.Z. Zhang, and J.E. Bowers. Electroabsorption modulator as a compact OTDM demultiplexing receiver. In *ECOC, paper 9.4.6*, 2000.
12. V. Kaman, A.J. Keating, S.Z. Zhang, and J.E. Bowers. Simultaneous OTDM demultiplexing and detection using an electroabsorption modulator. *IEEE Photonics Technology Letters*, 12:711–713, 2000.
13. V. Kaman, S. Yuan, O. Jerphagnon, R.J. Helkey, and J.E. Bowers. Comparison of wavelength-selective cross-connect architectures for reconfigurable all-optical networks. In *Photonics in Switching Conference, paper 14.1*, 2006.
14. V. Kaman, X. Zheng, S. Yuan, J. Klingshirn, C. Pusarla, R.J. Helkey, O. Jerphagnon, and J.E. Bowers. A 32×10 gbit/s DWDM metropolitan network demonstration using wavelength-selective photonic cross-connects and narrow-band EDFAs. *IEEE Photonics Technology Letters*, 17:1977–1979, 2005.
15. G. Karmous-Edwards. Today's optical network research infrastructures for E-Science applications. In *Optical Fiber Communications Conference (OFC/NFOEC), paper OWU3*, 2006.
16. K. Kato and Y. Tohmori. PLC hybrid integration technology and its application to photonic components. *IEEE Journal in Selected Topics on Quantum Electronics*, 6(1):4–13, 2000.
17. J. Kim, C.J. Nuzman, B. Kumar, D.F. Liewen, J.S. Kraus, A. Weiss, C.P. Lichtenwalner, A.R. Papazian, R.E. Frahm, N.R. Basavanahally, D.A. Ramsey, V.A. Aksyuk, F. Pardo, M.E. Simon, V. Lifton, H.B. Chan, M. Haueis, A. Gasparyan, H.R. Shea, S. Arney, C.A. Bolle, P.R. Kolodner, R. Ryf, D.T. Neilson, and J. V Gates. 1100×1100 port MEMS-based optical crossconnect with 4-dB maximum loss. *IEEE Photonics Technology Letters*, 15:1537–1539, 2003.
18. J. Kondis, B.A. Scott, A. Ranalli, and R. Lindquist. Liquid crystals in bulk optics-based DWDM optical switches and spectral equalizers. In *IEEE LEOS Annual Meeting*, 1; 292–293, 2001.
19. E. Lach and K. Schuh. Recent advances in ultrahigh bit rate ETDM transmission systems. *IEEE Journal of Lightwave Technology*, 24:4455–4467, 2006.
20. D.M. Marom, D.T. Neilson, D.S. Greywall, C.-S. Pai, N.R. Basavanahally, V.A. Aksyuk, D.O. Lopez, F. Pardo, M.E. Simon, Y. Low, P. Kolodner, and C. A. Bolle. Wavelength-selective $1 \times k$ switches using free-space optics and MEMS micromirrors: Theory, design, and implementation. *IEEE Journal of Lightwave Technology*, 23:1620–1630, 2005.
21. B. Mason, J.M. Geary, J.M. Freund, A. Ougazzaden, C. Lentz, K. Glogovsky, G. Przybylek, L. Peticolas, F. Walters, L. Reynolds, J. Boardman, T. Kercher, M. Rader, D. Monroe, and L. Ketelsen. 40 Gb/s photonic integrated receiver with -17 dBm sensitivity. In *Optical Fiber Communications Conference (OFC), paper FB10*, 2002.
22. B. Mason, A. Ougazzaden, C.W. Lentz, K.G. Glogovsky, C.L. Reynolds, G.J. Przybylek, R.E. Leibenguth, T.L. Kercher, J.W. Boardman, M.T. Rader, J.M. Geary, F.S. Walters, L.J. Peticolas, J.M. Freund, S.N.G. Chu, A. Sirenko, R.J. Jurchenko, M.S. Hybertsen, L.J.P. Ketelsen, and G. Raybon. 40-Gb/s tandem electroabsorption modulator. *IEEE Photonics Technology Letters*, 14(1):27–29, 2002.

23. S. Mechels, L. Muller, D. Morley, and D. Tillett. 1D MEMS-based wavelength switching subsystem. *IEEE Communications Magazine*, 41:88–94, 2003.
24. B. Mikkelsen, G. Raybon, R.J. Essiambre, K. Dreyer, Y. Su, J.E. Johnson, G. Shtengel, A. Bond, D.G. Moodie, and A.D. Ellis. 160 Gbit/s single-channel transmission over 300 km nonzero-dispersion fiber with semiconductor based transmitter and demultiplexer. In *ECOC, paper PD 2-3*, 1999.
25. M. Nakazawa, T. Yamamoto, and K.R. Tamura. 1.28 Tbit/s - 70 km OTDM transmission using third- and fourth-order simultaneous dispersion compensation with a phase modulator. In *ECOC, paper PD 2.6*, 2000.
26. M.J. O’Mahony, C. Politi, D. Klonidis, R. Nejabati, and D. Simeonidou. Future optical networks. *IEEE Journal of Lightwave Technology*, 24:4684–4696, 2006.
27. S. Pau, J. Yu, K. Kojima, N. Chand, and V. Swaminathan. 160-Gb/s all-optical MEMS time-slot switch for OTDM and WDM applications. *IEEE Photonics Technology Letters*, 13:1460–1462, 2001.
28. H.N. Poulsen, K.S. Jepsen, A.T. Clausen, A. Buxens, K.E. Stubkjaeri, R. Hess, M. Dulk, and G. Melchior. Fast optical signal processing in high bit rate OTDM systems. In *IEEE LEOS Annual Meeting*, volume 1, pages 67–68, 1998.
29. G. Raybon, P.J. Winzer, and C.R. Doerr. 1-Tb/s (10×107 Gb/s) electronically multiplexed optical signal generation and WDM transmission. *IEEE Journal of Lightwave Technology*, 25:233–238, 2007.
30. J.-K. Rhee, I. Tomkos, and M.-J. Li. A broadcast-and-select OADM optical network with dedicated optical-channel protection. *IEEE Journal of Lightwave Technology*, 21(1):25–31, 2003.
31. A.A.M. Saleh and J. M. Simmons. Evolution toward the next-generation core optical network. *IEEE Journal of Lightwave Technology*, 24:3303–3321, 2006.
32. R.E. Wagner, J.R. Igel, R. Whitman, M.D. Vaughn, A. B. Ruffin, and S. Bickham. Fiber-based broadband-access deployment in the United States. *IEEE Journal of Lightwave Technology*, 24:4526–4540, 2006.
33. H.G. Weber, R. Ludwig, S. Ferber, C. Schmidt-Langhorst, M. Kroh, V. Marembert, C. Boerner, and C. Schubert. Ultrahigh-speed OTDM-transmission technology. *IEEE Journal of Lightwave Technology*, 24:4616–4627, 2006.
34. S.J.B. Yoo. Optical packet and burst switching technologies for the future photonic internet. *IEEE Journal of Lightwave Technology*, 24:4468–4492, 2006.
35. M. Zirngibl. IRIS: Optical switching technologies for scalable data networks. In *Optical Fiber Communications Conference (OFC/NFOEC), paper OWP1*, 2006.

Index

- admissible solution, 268, 270, 271, 273, 274, 276
- aggregated demand, 257
- all-optical circuit grooming, 284
- all-optical edge grooming, 286
- all-optical grooming, 279
- all-optical wavelength grooming, 280
- amplifier, 124
- arbitrary alternate routing, 208

- bandwidth-blocking probability (BBP), 113
- bi-directional path networks, 69
- bi-directional ring networks, 69
- blocking probability, 193
 - end-to-end, 201
 - fiber route, 204, 206
 - g-link, 195, 197, 201, 204, 206, 208
 - path, 195, 197, 201, 208
 - segment, 204, 206
 - wavelink, 197, 199
 - active call, 207
 - passive call, 207
 - waveroute, 196, 197, 201, 204
- BXC, 254–256, 260, 261, 269, 271, 276

- call
 - active, 206
 - class, 198
 - multi-service, 198
 - passive, 206
 - service, 198
- carried load
 - fiber route, 205

- g-link, 200
 - active call, 208
 - passive call, 208
- path, 200
 - wavelink, 197, 201, 205
 - waveroute, 200, 205
- central office (CO), 107
- class- (n, k) call, 198
- complete lightpath cycle, 131
- complexity of Traffic Grooming, 57
- composite signal, 255
- connected lightpaths, 130
- control plane support
 - ASON, 21
 - mapping GMPLS, 26
 - call control, 31
 - hierarchical routing, 30
 - link resource manager, 28
 - mapping GMPLS, 27
 - model, 22, 26
 - resource discovery, 28
 - routing areas, 32
 - routing database, 28
 - subnetwork point, 26, 27
 - subnetwork point pool, 26, 27
 - UNI/NNI, 25
- GMPLS, 21
 - adaptation capability, 28
 - bi-directional LSP, 31
 - bundling, 28
 - call control, 31
 - exchange point, 25
 - explicit label, 31

- explicit route, 31
- forwarding adjacency, 28, 30
- generalized label, 27, 31
- hierarchical routing, 30
- hierarchy, 22
- label set, 31
- link management, 22, 29
- LSP, 22, 26
- mapping ASON, 26, 27
- MIB, 28
- model, 22, 26
- OPEX reduction, 21
- overlay/peer/augmented, 25
- resource discovery, 28
- routing protocols, 22
- RSVP-TE, 22
- signaling, 30
- signaling protocols, 22
- SRLG, 29
- support for VCAT/LCAS, 32
- switching capability, 27, 28
- TE link, 22, 26
- testbeds, 35
- upstream label, 31
- virtual link, 30
- grooming data plane
 - OTN, 22, 24
 - SONET/SDH, 22, 24
 - VCAT/LCAS, 24, 33, 34
- L1 VPN, 33
- PCE, 30, 34
- SLA/SLS, 33
- cost function, 261, 272
- cost per bit, 256
- crossbar switch, 17

- data centers, 253
- database backups, 253
- dedicated backup multiplexing, 152
- dedicated backup reservation, 150
- dedicated protection, 276
- demand time conflict reduction, 164
- dependable dynamic traffic grooming, 149
- deterministic, 254
- Digital Cross-Connect System (DCS), 15
- distribution
 - exponential, 111
 - Poisson, 111

- Dynamic Bandwidth Allocation (DBA), 51
- dynamic provisioning, 142
- dynamic restoration, 138
- dynamic traffic, 5

- electrical aggregation, 257
- electrical cross-connect, 254, 255
- electrical grooming, 254, 256, 257, 262, 272, 276
- elemental network topologies, 61
- End-to-end call loss probability, 197
- EXC, 254–257, 259, 260, 262, 266, 276

- Fiber Channel (FC), 108
- Fiber Connection (FICON), 108
- Fiber Cross-Connect (FXC), 12
- fixed-point equation, 197, 202

- g-link, 195, 196
 - carried load, 197
 - repeated, 209
 - support, 195
 - topology, 195
- general shared path protection, 144
- Generic Framing Procedure (GFP), 52, 108, 215
- grid networks, 69
- grooming, 254
- grooming and routing (GR), 123
- grooming lightpath, 257–259, 263, 264, 272, 273
- grooming link, 195
- grooming strategy, 256
- grooming switches, 15–19
- grooming, routing and wavelength assignment (GRWA), 124

- heuristics, 254
- hierarchical grooming
 - clustering, 79
 - general topology networks, 77
 - hub selection, 79
 - lightpath grooming, 83
 - logical topology, 81
 - performance results, 83
 - ring networks, 74
 - RWA, 82
 - star networks, 76

- torus networks, 76
- traffic routing, 81
- tree networks, 76
- hierarchical grooming algorithm, 79
- I/O ports, 256, 260
- Integer Linear Programming (ILP), 124
- inter-office (IOF), 107
- inverse multiplexing, 109
- Lagrangian relaxation, 124
- light tree, 18
- light-trail, 155, 165
 - dynamic, 168
 - static, 166
- lightpath, 2, 254–256, 260, 261, 264, 268–270, 272, 273
- lightpath cycle, 131
- lightpath sequence, 194
- Line Terminal Equipment (LTE), 59
- linear programming, 254
- Link Capacity Adjustment Scheme (LCAS), 51, 108
- link protection, 276
- link-based protection, 140
- logical topology, 2, 195
 - dynamic, 195
- management burden, 256
- Manhattan street networks, 69
- many-to-one traffic grooming, 227–232
 - exact approach, 229
 - heuristic approach, 231
- marginal demands, 259, 264
- matrix algebra, 263
- meta-heuristics, 254
- MG-OXC, 255, 260, 261
- mixed primary-backup grooming policy, 149
- monetary cost, 254
- Multi-Commodity Flow (MCF), 64
- multi-domain traffic grooming
 - conclusions, 251
 - introduction, 237
 - multi-domain standards, 238
 - open problems in multi-domain grooming, 245
 - research studies, 242
- multi-granularity switching, 255, 256
- multi-hop full grooming OXC, 17
- multi-hop partial grooming OXC, 18
- multi-layer, 255, 256, 262, 276
- Multi-Service Provisioning Platform (MSPP), 15
- multicast traffic grooming, 214–227
 - dynamic multicast traffic grooming, 222–227
 - enabling technology, 215
 - static multicast traffic grooming, 216–222
 - exact approaches, 216–220
 - heuristic approaches, 220–222
- multipoint traffic grooming, 213–233
 - types, 213
- network cost, 256, 261–263, 272, 276
- network design, 254
- network layer, 256, 261, 276
- node-edge incidence matrix, 126
- nominal capacity, 257
- NP*-Complete, 57, 112, 123
- objective function, 261
- OEO, 3
 - switching cost, 3, 60
- offered load
 - active call, 207
 - g-link, 200
 - passive call, 207
 - segment, 205
 - wavelink, 197, 201, 205
 - waveroute, 197
- optical burst switching (OBS), 287
- Optical Cross-Connect (OXC), 12, 114, 137
- optical grooming, 254, 256, 260–262, 268, 270, 273, 274, 276
- optical layer, 125
- optical node architectures, 9–19
- optical packet switching (OPS), 287
- optical switching, 13
- optical time-division multiplexing (OTDM), 284
 - principle, 285
- Optical Transport Network (OTN), 108
- optimisation techniques, 254
- opto-electronic, 256
- p*-Cycle protection, 141

- partial protection, 153
- partition, 269
- path protection, 276
- path restoration algorithm, 142
- path-based protection, 140
- performance model, 193
- permanent connection, 256
- permanent demands, 254
- physical layer, 125
- planning, 254
- Point-to-Point Protocol (PPP), 108
- power, 269
- pre-cross-connected shared path
 - protection, 145
- preplanned protection, 138
- protection design, 143
- provisioning, 254, 256, 273

- random traffic demands, 253
- random traffic grooming, 209
- random wavelength assignment, 209
- Reconfigurable OADM (ROADM), 280
 - multi-degree, 282
- reduced load approximation, 194, 196
- restricted sharing admission, 196, 198
- Routed Scheduled Band Group, 261, 270, 273, 274
- router, 255
- Routing and Wavelength Assignment (RWA), 3, 4

- scheduled demands, 253, 254, 276
- scheduled lightpath demands, 268
- scheduled services, 160
- SED, 254, 256, 257, 262–265, 267, 272, 273, 276
- segment, 204
- segregated primary-backup grooming
 - policy, 149
- self-routing fabric, 17
- service- k call, 198
- set-up date, 253, 254, 261, 262, 268
- shared backup multiplexing, 151
- shared backup reservation, 150
- shared bus, 16
- shared memory, 17
- shared protection, 276
- shelf, 269
- shortest path, 256, 262, 263, 268, 272

- single-hop grooming OXC, 17
- single-layer, 256
- SLD, 254, 256, 261, 262, 267–270, 273, 274, 276
- sliding scheduled traffic model, 162
- Soft Failure, 51
- solution space, 270
- SONET/SDH, 39
 - frame structures, *see* STS
 - layers, 40
 - Next Generation (NG-SONET/SDH), 47
 - rings
 - protection methods, 43
 - types, 43
 - standard rates, 41
 - Virtual Tributaries (VT), 43
- SONET/SDH ring grooming, 91–105
 - arbitrary traffic, 98–105
 - Bidirectional Line-Switched Ring (BLSR), 91
 - fixed traffic, 92–98
 - algorithms, 96–98
 - bounds, 93–96
 - hubbed rings, 95, 98
 - incremental traffic, 104
 - multi-ring topologies, 97
 - traffic types, 91
 - Unidirectional Path-Switched Ring (UPSR), 91
- source-edge incidence matrix, 126
- spider networks, 69
- star networks, 69
- state-of-the-art, 256
- static traffic grooming with min-max
 - objective, 71
- statistical multiplexing gain (SMG), 12
- STM, 41
- storage-area network, 108
- STS, 41
- subgraph-based routing, 140
- subgraph-based routing strategy, 142
- subwavelength demands, 3, 253–256
- survivability, 138
- switch ports, 254
- switching fabric, 255, 256
- switching granularity, 253

- tear-down date, 253, 254, 261, 262, 268, 276
- time diagram, 257, 272
- time-disjoint, 254, 276
- topology, 254, 257, 260–263
- traffic grooming, 159
 - dynamic, 193
 - multi-hop, 194
 - policy, 206
 - single-hop, 194
 - static, 193
- traffic pattern, 111
- transit traffic, 256
- transponder, 12, 124
- transponders, 254
- tree networks, 71
- tuple, 254, 261

- uncertainty, 254, 276
- unidirectional path networks, 62
- unidirectional ring networks, 69

- ventilation system, 269
- Virtual Concatenation (VC/VCAT), 48, 108, 109
- Virtual concatenation group (VCG), 48
- virtual topology, 2

- waveband, 253–256, 260, 261, 270, 271
- waveband cross-connect, 254, 255
- waveband switching, 261
- waveband-switching connection, 260, 261, 267, 268, 271
- wavelength, 255
- wavelength assignment, 260
- wavelength assignment (WA), 123
- wavelength capacity, 125
- wavelength continuity, 126
- wavelength continuity constraint, 260
- wavelength conversion, 204, 260
- wavelength cross-connect, 254, 255
- Wavelength Interchanging Cross-Connect (WIXC), 13
- Wavelength Selective Cross-Connect (WSXC), 13
- Wavelength-Band Cross-Connect (WBXC), 12
- wavelink, 195, 196, 198
 - repeated, 204, 209
- wavelink blocking model, 198
- waveroute, 195, 196, 198, 204
- WDM grooming, 193
- WXC, 254–256, 259–262, 266, 269, 271, 276

Part II

Techniques

Traffic Grooming in SONET/SDH Rings

Raza Ul-Mustafa

7.1 Introduction

This chapter deals with the problem of traffic grooming on WDM ring networks employing SONET equipment. The chapter addresses the problems of dimensioning and provisioning such networks in order to support static traffic. We briefly review the pertinent features of SONET rings in the next paragraph; more details are available in Chapter 4 and the references cited there in.

SONET rings, which are very widely used in Metro networks, are of two types:

- Unidirectional Path-Switched Ring (UPSR), Fig. 7.1(a), and
- Bidirectional Line-Switched Ring (BLSR), Fig. 7.1(b).

UPSR rings use two fibers for transmission in two opposite directions, where one fiber is the working fiber and the other is the protection fiber. The signal from a source is transmitted on both fibers, and the receiver monitors both of them, and then selects the better of the two signals. BLSR rings use either two or four fibers, hence called BLSR/2 and BLSR/4, respectively. In the BLSR/4 ring (shown in Fig. 7.1(b)), two fibers are used as working fibers, and the other two are used for protection. However, the working traffic can be carried in either of the two directions (typically using the shortest path), and the protection traffic is transmitted in the opposite direction, only in the case of failure. BLSR/2 is similar in topology to UPSR, but is similar in operation to BLSR/4. The use of two fibers in BLSR/2 requires that the capacity of each fiber be split between working and protection bandwidth.

SONET rings in metro networks are typically provisioned to support static traffic. Under static traffic, a fixed matrix of traffic demands is available, i.e., all the demands are known a priori. Given a physical topology, the task is to design and provision the network such that the network cost is minimized. The traffic matrix itself can be further classified into one of the following types:

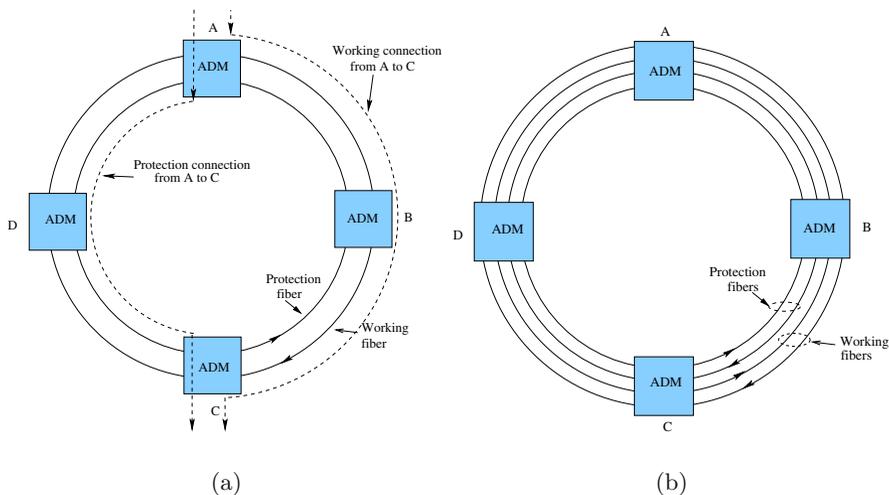


Fig. 7.1. SONET rings: (a) UPSR; (b) BLSR/4

1. *Uniform* traffic demands, where each pair of nodes in the network exchanges the same level of traffic units,
2. *Distance-dependent* traffic demands, where the number of traffic units between a node pair depends on the distance between them (i.e., a node shares more traffic with the neighboring nodes than the nodes farther from it), and
3. *Non-uniform* traffic demands, which is the more general case, where the number of traffic units between node pairs is arbitrary.

We will refer to the two special cases of uniform and distance-dependent traffic patterns, as the *fixed traffic pattern*, and to the case of non-uniform traffic demands as *arbitrary traffic pattern*.

Some of the earliest studies in the area of traffic grooming on ring architecture with static traffic patterns were conducted by Simmons, Goldstein, and Saleh [16, 17, 18], Gerstel, Lin and Sasaki [9, 10], Chiu and Modiano [4, 5], and Zhang and Qiao [24, 25]. All such studies advocated that, to reduce the network cost, more emphasis should be placed on minimizing the number of higher layer components than the number of wavelength channels. However, this is a hard problem and was proven to be NP-Complete by Chiu and Modiano [5] and Li et al. [13] by reduction from the bin packing problem (see Chapter 5 for a detailed discussion on the complexity of the traffic grooming problem). Researchers addressed the problem by providing bounds, heuristics, and exact solutions. Table 7.1 provides a classification of some of the related work in this area.

Table 7.1. Related work in the area of traffic grooming on ring architecture with static traffic patterns

Traffic Pattern	Solution Approach	References
Fixed Traffic	Bounds	[1, 3, 4, 5, 10, 16, 17, 24]
	Heuristics	[3, 4, 5, 16, 17, 24, 25]
	ILPs	[11]
Arbitrary Traffic	Bounds	[8, 13, 21, 22, 23, 25]
	Heuristics	[8, 13, 19, 21, 22, 23, 25, 26]
	ILPs	[8, 11, 22]

7.2 Fixed Traffic Patterns

In this section, we will review the literature related to the design of WDM networks for traffic grooming on ring topologies using fixed traffic patterns.

7.2.1 Bounds for General Ring Topologies

Simmons, Goldstein, and Saleh provided some initial work in [16] and later refined it in [17]. In both references, they considered a ring topology with fixed traffic pattern. Specifically, in [17], the authors considered uniform all-to-all traffic and distance-dependent traffic between nodes connected as a ring topology. Uniform all-to-all traffic is defined as a traffic scenario whereby a single unit of traffic is sent from each node on a ring to every other node on that ring, while distance-dependent traffic is defined as a traffic scenario whereby the number of traffic units sent between nodes increases as the inter-nodal distance decreases. The authors provided upper bounds on the number of SONET Add/Drop Multiplexers (ADMs) and also provided procedures to obtain these bounds. If r represents the number of units of traffic between a source-destination pair and G represents the total number of traffic units that a single wavelength can accommodate,¹ then Equation (7.1) expresses the estimate given in [17] of the required number of ADMs per wavelength for a scenario where each of the N nodes² in a ring is transmitting $1/G$ units of traffic to every other node in that ring. Similarly, Equation (7.2) provides an estimate for the total number of required ADMs for the scenario where each node transmits a full wavelength to each other node, i.e., $r = G$.

$$r = \frac{1}{G} : \text{Number of ADMs per wavelength} = \frac{4N}{N+1} \quad (7.1)$$

$$r = G : \text{Number of ADMs per wavelength} = \frac{N(N-1)}{2} \quad (7.2)$$

¹ The terms *grooming factor* or *grooming ratio* are used interchangeably to refer to G in the rest of this chapter.

² N was assumed to be *odd*.

They also introduced the interesting concept of a *super node*, where nodes are grouped into distinct super nodes such that the amount of traffic between super nodes occupies a full wavelength. The traffic between super nodes can then be routed using procedures developed for full wavelength traffic. The nodes, however, were grouped into super nodes using approximations, and hence the overall solution is an approximate one. In [17] the authors indicated that, to reduce the number of ADMs, more connections need to be routed optically at each node and have therefore quantified the benefit of using wavelength add-drops by measuring the through-to-total ratio of connections at each node in the network. In order to analyze the distance-dependent traffic scenario, they considered a traffic pattern on a bidirectional ring with an odd number N of nodes for which the amount of traffic between the most distant nodes is one unit, and the traffic demand increases by one unit as the inter-nodal distance decreases by one link. They estimated the number of units of traffic sent by each node, in each direction to be $(N^2 - 1)/8$, and the number of traffic units on each link to be $(N^2 - 1)(N + 3)/48$. The minimum number of required ADMs, for $G = 1$, to support this traffic was then estimated as:

$$G = 1 : \text{Minimum number of ADMs} = \frac{N(N^2 - 1)}{8} \quad (7.3)$$

Finally, another important contribution of this work was the realization that minimizing the number of ADMs will not only reduce the cost of the network, but also in some cases, such as IP over WDM, the ADM savings may imply a substantial simplification in packet routing.

7.2.2 Bounds for UPSR and BLSR/2 Rings

In [4], Chiu and Modiano considered all-to-all uniform traffic on WDM unidirectional ring networks. In [5], they extended their work by considering distance dependent-traffic also. One major contribution of this work is the introduction of upper and lower bounds for the required number of wavelengths and ADMs. Furthermore, they proposed heuristic solutions with a performance that is close to the derived bounds. Using the same definitions for r , G and N above, the authors in [5] give the following loose lower bound on the number of required ADMs:

$$\text{Number of ADMs} \geq \left\lceil \frac{(N - 1)r}{G} \right\rceil N \quad (7.4)$$

They also provided two upper bounds on the minimum number of required ADMs, which are shown in Equations (7.5) and (7.6):

$$\text{Minimum number of ADMs} \leq \left\lceil \frac{N(N - 1)r}{2G} \right\rceil N \quad (7.5)$$

$$\text{Minimum number of ADMs} \leq \left\lceil \frac{r}{G} \right\rceil N(N - 1) \quad (7.6)$$

They also proposed two heuristic approaches to accommodate all-to-all uniform traffic while minimizing the required number of ADMs. The first heuristic attempts to maximize the number of nodes that require fewer ADMs. For example, since a node uses k ADMs if it transmits/receives on k wavelengths, then if M_k represents the number of nodes with k ADMs, the heuristic first tries to maximize M_1 , then M_2 , then M_3 , and so on. The second heuristic assigns nodes to wavelengths by attempting to pack the wavelengths efficiently. All of the N nodes are first divided into $N/\lfloor\sqrt{G}\rfloor$ groups. The cross traffic of each pair in a group is then assigned to a wavelength. This is possible because, by design, the cross traffic between two groups of size $\lfloor\sqrt{G}\rfloor$ is less than G . The traffic within each group is first accommodated on existing wavelengths, and then on additional wavelengths if needed. Besides all-to-all-uniform traffic, the authors also discussed the distance-dependent traffic and derived the upper and lower bounds on the number of ADMs shown in Equations (7.7) and (7.8) below, respectively.

$$\text{Number of ADMs} \leq N \left\lceil \frac{N(N^2 - 1)}{8G} \right\rceil \quad (7.7)$$

$$\text{Number of ADMs} \geq \left\lceil \frac{N(N^2 - 1)}{4G} \right\rceil \quad (7.8)$$

Finally, one of the results mentioned in [5] suggests that the number of wavelengths and ADMs cannot always be minimized simultaneously. However, the authors conjectured that for the special case where each source-destination pair has a single unit of traffic between them, the minimum number of ADMs could be achieved with the minimum number of wavelengths. Later, in [1] Bermond and Coudert proved this conjecture to be true for few values of G (e.g., $G = 3$), while proving it to be false for many other values of G (e.g., $G = 7$).

Reference [10] focused on obtaining bounds on the number of SONET ADMs when nodes are placed on either a UPSR or a BLSR. The authors, however, restricted their analysis to uniform traffic between the nodes. They considered single-hub UPSRs, and assumed that low speed traffic can be cross-connected at nodes. Equation (7.9) shows the lower bound on the number of required ADMs when a UPSR ring is considered (with uniform traffic), while Equation (7.10) gives the lower bound when a single-hub UPSR ring is considered.

$$\text{UPSR: Number of ADMs} \geq \max \left\{ \left\lceil \frac{2N(N-1)r}{G+r} \right\rceil, N \right\} \quad (7.9)$$

$$\text{UPSR, hub: Number of ADMs} \geq (N-1) \left\lceil \frac{(N-1)r}{G} \right\rceil + (N-1)k + \lambda_f \quad (7.10)$$

In Equation (8.10), $k = \lfloor (N - 1)r/G \rfloor$ is the number of wavelengths dedicated to each non-hub node, and λ_f is the number of wavelengths used to carry any leftover traffic for non-hub nodes that was not accommodated by the k wavelengths. After comparing their computed lower bound for single-hub UPSR with the one given in [5], which assumes that traffic streams cannot be cross-connected, the authors concluded that using cross-connection usually results in using fewer ADMs. However, the drawback of using a single-hub architecture is the increase in the number of wavelength channels. The authors also demonstrated that although spatial reuse of the bandwidth of BLSR/2 ring networks makes it less expensive to implement than UPSR, in a few cases (for which $r \leq G/2$) UPSR cannot be more expensive than BLSR/2 in terms of the number of ADMs needed to accommodate the traffic. Finally, the authors considered a network architecture with two different line speeds, namely, OC-12 ($G = 4$) and OC-48 ($G = 16$). They plotted the lower bounds on the cost using Equation (7.9), while using a cost-factor of 2.5 for OC-48 over OC-12. The conclusions were that the cost of using OC-48 is often less than the cost of using OC-12.

In reference [3], the authors computed lower bounds on the number of wavelengths in a BLSR network under static all-to-all uniform traffic. They also computed lower bounds on the number of ADMs in unswitched UPSR rings for uniform all-to-all traffic, which are given by Equation (7.11):

$$\begin{aligned} \text{Number of ADMs} &\geq \max \left\{ \left\lceil \frac{N(N-1)(n+f)}{n(n-1)(1-f) + 2\lfloor G/r \rfloor f} \right\rceil, N \right\} \\ n &= \left\lceil r + \frac{\sqrt{r^2 + 8rG}}{2r} \right\rceil \\ f &= \begin{cases} 1, & \text{if } \frac{2\lfloor G/r \rfloor}{n-1} - n \geq 1 \\ 0, & \text{otherwise} \end{cases} \end{aligned} \quad (7.11)$$

For the UPSR case, the authors also presented a heuristic approach that first constructs full circles, and later grooms these circles onto wavelengths. They focused on balancing the ADM utilization at every node and not splitting a traffic stream into more than one wavelength. The circles are constructed in a manner similar to that of [25]. To assign the circles onto wavelengths a multi-phase approach is used. In phase 1, nodes are divided into clusters such that the traffic between all the nodes of a cluster could be accommodated by one wavelength without splitting any stream. After that, the traffic between the nodes from two different clusters is groomed, provided it results in an effective use of ADMs. In phase 2, more circles are groomed onto existing wavelengths provided that a wavelength has enough capacity and the addition of an ADM does not decrease the ‘‘circles per ADM’’ factor (the factor captures the efficiency of ADM utilization). In phase 3, any remaining circles are groomed by adding such circles to a wavelength that share one or more end points.

7.2.3 Algorithms Based on Graph Theory

A number of studies have used graph theoretic approaches to study the problem of SONET ring network provisioning under the case of static traffic. The authors in reference [1] used tools from graph theory and design theory to address the traffic grooming problem in unidirectional WDM ring networks. They also considered uniform all-to-all traffic demands among the nodes of the network. They showed that the problem of minimizing the number of ADMs for uniform all-to-all traffic demands can be expressed by partitioning the edges of a complete graph of N vertices into W subgraphs, where W is the number of wavelengths such that the total number of vertices in each subgraph is minimized. Each subgraph in this case represents a wavelength, while the vertices in each subgraph correspond to ADMs. The authors pointed out that this problem is similar to design theory, such that an $(N, k, 1)$ -design is nothing but a partition of the edges of a graph into subgraphs isomorphic to blocks in this theory. The classical equivalent definition is as follows: *given a set of N elements, find a set of blocks such that each block contains k elements and each pair of elements appears in exactly one block.* Thereafter, using techniques developed in design theory, the authors computed the number of ADMs required for many possible combinations of the grooming factor and the number of nodes in the network.

7.2.4 Algorithms Based on Meta-Heuristics

Some researchers have used meta-heuristics to solve the traffic grooming problem on WDM ring networks. In reference [6], the authors used Simulated Annealing as one part of the solution approach. They considered WDM ring networks with uniform traffic and presented two different approaches. In one approach, the traffic needs to be routed through a hub node in the network. They defined this approach as a *multihop* approach, because each traffic demand delivery needed to employ more than one lightpath hop. Also, they consider the case in which traffic can be routed without going through any hub node and called it *single-hop* approach. For the multihop approach, they first place ADMs on the nodes for each request and then try to re-assign the traffic such that the maximum number of ADMs and wavelengths can be reduced. For the single-hop approach, they employed a two-phase methodology. In the first phase, they constructed the virtual circles, while in the second phase they used Simulated Annealing to group the circles into different wavelengths. Based on simulation results, they argued that it was beneficial in terms of the number of ADMs to use the single-hop approach while employing simulated annealing for small grooming ratios. However, for large grooming ratios and large networks, the multihop approach could lead to better savings in terms of the number of ADMs.

7.2.5 Algorithms for Multi-Ring Topologies

The authors in reference [15] explored an interesting variant of the ring topology. They considered a topology in which many rings are connected to each other forming a WDM multi-ring network. They classified each traffic connection as either *intra-connection* or *inter-connection*. For an intra-connection, the source and destination are located on the same ring, while for an inter-connection the source and destination are located on different rings. Four virtual topologies are formed depending on how inter-connection traffic is transmitted and also how those connections are grouped onto a wavelength. The four topologies are referred to as: independent virtual topology (IVT), separated virtual topology (SVT), mixed virtual topology (MVT), and partial mixed virtual topology (PMVT).

- In IVT, an inter-connection is cut at the inter-ring connection node, and therefore all traffic is considered as intra-connection. Accordingly, at an inter-ring connection node every wavelength needs to have an ADM.
- In SVT, a connection is strictly classified as intra-connection or inter-connection, and a wavelength carries only one type of connection. An inter-ring connection node thus needs ADMs only for the inter-connection wavelengths.
- In MVT, a wavelength could carry inter-connection traffic, intra-connection traffic, or both.
- In PMVT, some intra-connections and inter-connections are groomed together onto a wavelength to form an MVT, while other connections are groomed separately as in SVT.

Finally, to groom the traffic for each type of virtual topology, the authors proposed four heuristics, the independent traffic grooming (ITG), the separated traffic grooming (STG), the mixed traffic grooming (MTG), and the partially mixed traffic grooming (PMTG), for IVT, SVT, MVT and PMVT, respectively. Using simulation, they showed that ITG requires the minimum number of wavelengths but uses the most number of ADMs at an inter-ring connection node, MTG requires the maximum number of wavelengths, while STG and PMTG require the minimum number of ADMs.

7.3 Arbitrary Traffic Patterns

In this section, we will review the literature related to the design of WDM networks for traffic grooming on ring topologies in order to support arbitrary traffic between the different node pairs.

7.3.1 Single-Hub SONET Rings

Li et al., in reference [13] extended the work done in reference [10] by focusing only on single-hub SONET/WDM rings, while considering uniform and non-

uniform traffic. First, they proved that, given a set of traffic demands, the single-hub BLSR/2 costs no more than the single-hub UPSR under any traffic pattern³. They then showed that optimal traffic grooming can be confined to a narrow subset of valid grooming strategies, and referred to these as *canonical groomings*. For uniform traffic, they computed the required number of ADMs in the working fiber of a UPSR ring, $F(G, r, N)$, which is given by Equation (7.12). The total number of required ADMs for a UPSR ring is then $2F(G, r, N)$, and for a BLSR is $F(G/2, r, N)$.

$$F(G, r, N) = N \left\lceil \frac{r}{G} \right\rceil + N \left\lfloor \frac{r}{G} \right\rfloor + \left\lceil \frac{N}{\lfloor G/(r \bmod G) \rfloor} \right\rceil \quad (7.12)$$

For non-uniform traffic, the authors showed that grooming on a single-hub BLSR/2 could be solved using bin-packing approaches. They proved that using a First-Fit-Decreasing (FFD) algorithm [7] to assign non-uniform integer traffic units to ADMs produces a 10/9 approximation ratio. A typical first-fit-decreasing algorithm first sorts the input objects in decreasing order, and then assigns the objects (sequentially) to a first-fit bin. Therefore, the authors first sorted the residual traffic demands, $r_i \bmod G$, at all non-hub nodes decreasingly and then assigned the demands to the first ADM with sufficient capacity. Note that the remaining traffic demands at each node occupy $\lfloor r_i/G \rfloor$ full wavelengths, and each wavelength in the case of the single-hub architecture requires two ADMs. Finally, for non-uniform integer traffic demands and grooming factors of 2, 4, and 8, the authors also presented optimal algorithms by solving an integer bin packing problem. It turns out that in a few cases the solution strategy is similar (but not exactly the same) to that of FFD.

7.3.2 SONET BLSR

In reference [21], the authors considered SONET/WDM BLSRs, while taking into account non-uniform traffic. They assumed each traffic stream to consist of unitary traffic demand. However, each traffic stream can occur between any pair of arbitrary nodes (as opposed to all-to-all or one-to-all traffic). Therefore, such a setting permits non-uniform traffic. They extended the work presented in reference [20] by considering two versions of the minimum ADM cost problem. In the first version, named *arc-version*, each traffic stream has its predetermined routing, such as shortest path routing. Therefore, the two (working) fiber rings⁴ are handled separately, such that in each fiber ring, a traffic stream is represented as a (directed) circular arc. For the second

³ Compare this result with [10] in which they showed that in a few cases UPSR does a little better than (non-hub) BLSR/2. These results do not contradict each other, as the conclusions in [13] are for single-hub rings only, while those of [10] are for non-hub rings.

⁴ In the case of a BLSR/2 ring, each fiber can be regarded as two virtual fibers, one used for working traffic and the other used for protection.

version, named *chord-version*, they assumed that each traffic stream is full-duplex with symmetric and unitary demands, which must be routed along the same path but in the opposite direction. Thus, the two (working) fiber rings are treated as one (undirected), and each traffic stream is treated as a (undirected) chord. In terms of solution approaches, they presented some lower bounds, and also presented a set of heuristics. To express these bounds, let A represent the set of input arcs, and let $\delta_A(i)$ and $\tau_A(i)$ represent the total number of arcs in A that originate (terminate) from (to) node i , respectively. A lower bound on total number of required ADMs is then given by:

$$\text{Number of ADMs} \geq \sum_{i=1}^N \left[\max \left(\delta_A(i), \frac{\tau_A(i)}{G} \right) \right] \quad (7.13)$$

To present a lower bound on the number of ADMs for the chord-version, let B represent the input set of chords and $\theta_B(i)$ represent the total number of chords in B that contain node i as one endpoint. The lower bound on ADMs is then given by:

$$\text{Number of ADMs} \geq \sum_{i=1}^N \left[\frac{\theta_B(i)}{2G} \right] \quad (7.14)$$

For the heuristic approach, the basic idea is to first construct *primitive rings* from the unitary traffic demands and then partition them into a number of groups such that each group consists of a set of primitive rings and the size of each set does not exceed the grooming factor G . Given a set of primitive rings, they propose the following algorithm, named iterative matching, for constructing groups. Let m represent the total number of primitive rings, $m \bmod G = 0$ (which could be achieved by adding dummy rings) and G is a power of 2. Let Π_0 be the original sets. The i th iteration starts with Π_{i-1} , a 2^{i-1} -grouping of Π_0 , and finds a maximum-weighted perfect matching of Π_i . Then, for each edge in the obtained matching, the two sets incident to the edge are merged. Thus, the i th iteration outputs a 2^i -grouping of Π_0 , denoted by Π_i . The authors also extended the iterative matching algorithm for the case when G is not a power of 2. This work, however, did not include any performance study of the heuristics.

7.3.3 Two-Phase Heuristic Approaches

In [25], the authors extended their earlier work in [24] and proposed heuristic approaches to solve the traffic grooming problem on unidirectional and bidirectional rings, while considering both uniform and non-uniform traffic. The main idea of the heuristics is to follow a two-phase approach. In the first phase, *circles* are constructed using traffic streams, and in the second phase those circles are groomed onto the set of the available wavelengths. Three heuristics were presented to construct circles for uniform traffic, and one heuristic for the

construction of circles for non-uniform traffic was introduced. In addition, a heuristic for circle grooming was provided. The heuristic for the construction of circles using non-uniform traffic may be summarized as follows.

1. Construct circles with the traffic streams having the longest stride.
2. For each of the remaining traffic streams: accommodate it on an already constructed circle if it does not overlap with any other traffic stream present on that circle *and* it also shares at least one end node with any of the present traffic streams on that circle; otherwise, add this traffic stream to a GapMaker list.
3. For each traffic stream in the GapMaker list: try to accommodate the traffic stream on already constructed circles if the restrictions mentioned in Step 2 are met; otherwise, create a new circle with this traffic stream.

The heuristic presented for grooming of circles onto wavelengths first determines the number of circles, m_w , to be groomed on wavelength w . After that, for each wavelength w the following two steps are executed:

1. Find the circle which has the maximum number of ADMs over all existing circles and groom it onto wavelength w .
2. Repeat $(m_w - 1)$ times: groom such a circle onto wavelength w , which when groomed onto w , results in a minimum number of additional ADMs.

Another two-phase approach that addresses the traffic grooming problem on unidirectional and bidirectional rings is presented in [19]. For unidirectional rings, the authors first mapped the ring onto an extended linear topology. They defined a collection of traffic streams that do not share any physical link with each other as a *string*. Then, in the first step of their two-phase approach, they allocated the input traffic onto the extended linear topology while minimizing the number of the strings. In the second step, they employed a grouping algorithm to combine the strings into wavelengths while minimizing the total number of ADMs. In Fig. 7.2, a simple five-node unidirectional ring and its mapping onto a linear topology is depicted. Nodes 6, 7, 8, and 9 (also represented as 1', 2', 3', and 4', respectively) are the added dummy nodes, which correspond to nodes 1, 2, 3, and 4, respectively. Traffic sourced at a node and destined to a lower indexed node will now terminate at the corresponding added dummy node. Once all the traffic is mapped from a unidirectional ring onto a linear topology, the solutions developed for the linear topology below will be applicable to the unidirectional ring too. Figure 7.3 shows the output of the algorithm that minimizes the number of strings, when a sample input and a linear topology of 5 nodes is used. The traffic streams are shown in Fig. 7.3(a). Each segment corresponds to a single traffic unit. Figure 7.3(b) shows the traffic streams after sorting. Figure 7.3(c) shows the output of the string minimization algorithm after efficiently combining the non-overlapping traffic streams (as strings). For each string, the circles at nodes represent the streams with common end-points alluding that an ADM could be shared between these streams. The key idea of the grouping algorithm is to group together, into a

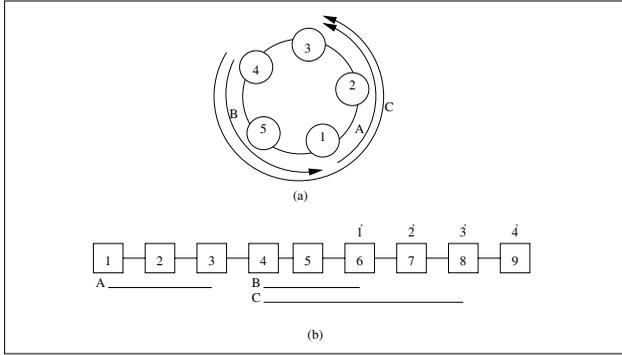


Fig. 7.2. Mapping a unidirectional ring into a linear topology; nodes 6 to 9 are the added dummy nodes corresponding to nodes 1 to 4, respectively

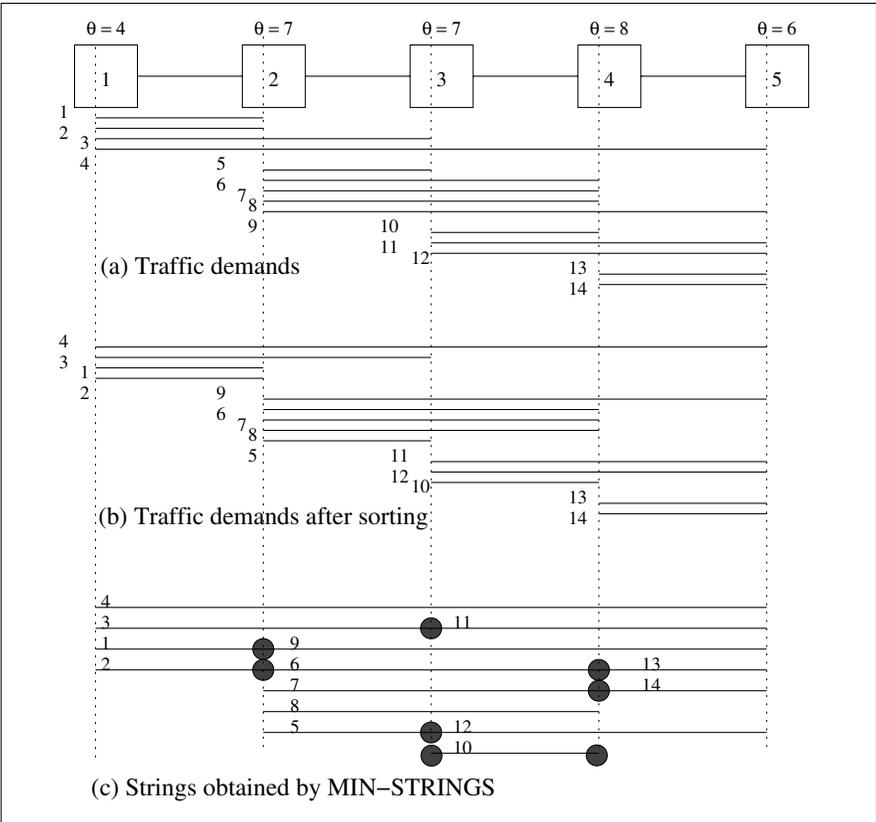


Fig. 7.3. Packing the traffic demands while minimizing the number of strings

wavelength, such strings that share maximum number of common end-points. This leads to sharing of ADMs, resulting in fewer ADMs for that wavelength.

For the bidirectional ring architecture, the authors developed techniques that map a bidirectional ring onto a set of unidirectional rings. They also proposed some traffic routing techniques that reduce the number of ADMs needed to accommodate the input traffic matrix. The time complexity of their techniques is shown to be at least an order of N lower than the other proposed two-phase approaches.

7.3.4 Meta-Heuristic Approaches

Meta-heuristics were also used to solve the SONET ring provisioning problem under arbitrary traffic patterns. For example, in reference [21] the authors presented a Genetic Algorithm-based approach for unidirectional SONET/WDM rings with arbitrary and asymmetric traffic. They chose an interesting representation of the *chromosome* by making use of an order-based approach, which essentially determines the order in which each traffic stream will be considered for grooming. Given a chromosome, they decoded the chromosome using a greedy heuristic, which follows a first-fit bin packing algorithm style. New offspring were generated using a $(\mu + \lambda)$ strategy, in which μ parents are used to generate λ children and all $(\mu + \lambda)$ solutions compete for the next generation.

7.3.5 Linear Programming Methods

Linear programming has also been employed to determine the optimal solution for the traffic grooming problem in SONET ring networks. In reference [22], the authors developed an integer linear programming (ILP) formulation for bidirectional WDM rings allowing arbitrary traffic between node pairs. They also presented approximate solutions using Simulated Annealing and a greedy approach. In reference [11], the authors also developed an ILP to minimize the total number of ADMs on a WDM ring network. Then, the complexity of the problem was reduced by relaxing a few integer variables to real variables but without compromising the solution quality for uniform traffic only. An extension to incorporate non-uniform traffic, using integer variables, was also provided. Finally, the authors also illustrated how this formulation can be extended to include dynamic traffic. Results were provided using fixed traffic only.

In [8], the authors formulated an ILP for the traffic grooming problem on unidirectional rings. The objective of the optimization problem was to minimize electronic routing. An interesting contribution of this work was the formulation of a sequence of bounds, where each successive bound in the sequence was tighter, or at least as good as the previous one. Each successive bound, however, involved more computation. The underlying principle in obtaining bounds was to decompose the ring into many sets of nodes, such

that in each set nodes were arranged in a path. A locally optimal solution was then employed within that set.

Recently, the use of SONET equipment equipped with transport blades operating at different line rates was suggested. The reason for doing this is to reduce the cost of implementation, since the cost of transport blades increases with the line rate. The optimal design of SONET rings with different line rates was studied using linear programming in a number of papers. In [12], an ILP was formulated to solve this problem. However, in [14] it was realized that an exact non-linear formulation can be decomposed into a number of smaller ILP formulations based on the fact that the identities of wavelengths being used at different rates are not as important as their number. Therefore, the ILP formulation corresponded to the number of combinations of wavelengths operating at different rates, which is significantly smaller than the number of wavelength permutations. This solution technique resulted in a significant saving in time and space.

7.3.6 Incremental Traffic

The following two studies considered incremental traffic patterns, and thus do not strictly fall under the category of static traffic. However, in both cases, once the initial static traffic has been accommodated, all of the new traffic are known at the same time and are treated collectively (as opposed to traffic arrival and accommodation separately). In [2], the authors defined a set of traffic matrices that reflects the changes in the traffic over a specified time period. The task was then to minimize the number of electronic ADMs needed to support any traffic matrix in this set; hence essentially supporting the traffic which dynamically changes within the given set of matrices. They considered a unidirectional ring. They also assumed that the traffic matrices are t -allowable, i.e., no node is sourcing more than t units of traffic at any time, where t is a specified constant. They then started out by placing an ADM on each wavelength at every node, and provided heuristics that remove the maximum possible number of ADMs, while supporting every t -allowable traffic matrix. They mentioned that their heuristic approaches can save up to 27% of the number of ADMs.

Similarly, researchers in [26] addressed the traffic grooming problem on unidirectional rings with incremental traffic using meta-heuristics. The authors assumed that they start with a given static traffic matrix and the network has been provisioned for that traffic matrix. Some time later, another traffic matrix is introduced. The objective is to accommodate the new traffic requests by disrupting as few current connections as possible. Two flavors of the problem are considered: best fit, in which they try to place as much new traffic as possible without adding any additional SONET ADMS, and full fit, in which they try to accommodate all the new traffic by adding minimum number of additional SONET ADMs. They presented an ILP, a

greedy heuristic, and a Tabu Search approach to solve the best fit case. They also presented an ILP and a Tabu Search based technique for the full fit case.

References

1. J.C. Bermond, D. Coudert (2003). Traffic grooming in unidirectional WDM ring networks: The all-to-all unitary case. In: *Proceedings of Optical Network Design and Modeling*, 2:1135–1153.
2. R. Berry, E. Modiano (2000). Reducing electronic multiplexing costs in SONET/WDM rings with dynamically changing traffic. *IEEE J. Selected Areas in Communications*, 18:1961–1971.
3. A.R. Billah, B. Wang, A.A. Awwal (2002). Effective traffic grooming in WDM rings. In: *Proceedings of IEEE GLOBECOM*, 3:2726–2730.
4. A. Chiu, E. Modiano (1998). Reducing electronic multiplexing costs in unidirectional SONET/WDM ring networks via efficient traffic grooming. In: *Proceedings of IEEE GLOBECOM*, 1:322–327.
5. A. Chiu, E. Modiano (2000). Traffic grooming algorithms for reducing electronic multiplexing costs in WDM ring networks. *J. Lightwave Technology*, 18:2–12.
6. W. Cho, J. Wang, B. Mukherjee (2001). Improved approaches for cost-effective traffic grooming in WDM ring networks: Uniform-traffic case. *J. Photonic Network Communications*, 3:245–254.
7. E.G. Coffman, M.R. Garey, D.S. Johnson (1997). Approximation algorithms for bin packing – A survey. In: *Approximation Algorithms for NP-Hard Problems*. PWS Publishing Company, Boston, pp. 46–93.
8. R. Dutta, G.N. Rouskas (2002). On optimal traffic grooming in WDM rings. *IEEE J. Selected Areas in Communications*, 20:110–121.
9. O. Gerstel, R. Ramaswami, G. Sasaki (1998). Cost-effective traffic grooming in WDM rings. In: *Proceedings of IEEE INFOCOM*, 1:69–77.
10. O. Gerstel, P. Lin, G. Sasaki (1999). Combined WDM and SONET design. In: *Proceedings of IEEE INFOCOM*, 2:734–743.
11. J.Q. Hu (2002). Traffic grooming in wavelength division multiplexing ring networks: A linear programming solution. *J. Optical Networking*, 1:397–408.
12. A. Jarray, B. Jaumard (2005). Exact ILP solution for the grooming problem in WDM ring networks. In: *Proceedings of IEEE ICC 2005*.
13. X.Y. Li, L.W. Liu, P.J. Wan, O. Frieder (2000). Practical traffic grooming scheme for single-hub SONET/WDM rings. In: *Proceedings of 25th IEEE Conference on Local Computer Networks*, pp. 556–564.
14. H. Liu, F. Tobagi (2004). A novel efficient technique for traffic grooming in WDM SONET with multiple line speeds. In: *Proceedings of IEEE ICC 2004*.
15. S.S. Roh, W.H. So, Y.C. Kim (2001). Design and performance evaluation of traffic grooming algorithms in WDM multi-ring networks. *J. Photonic Network Communications*, 3:335–348.
16. J.M. Simmons, E.L. Goldstein, A. Saleh (1998). On the value of wavelength-add/drop in WDM rings with uniform traffic. In: *Proceedings of Optical Fiber Communication Conference*, pp. 361–362.
17. J.M. Simmons, E.L. Goldstein, A. Saleh (1999). Quantifying the benefit of wavelength add-drop in WDM rings with distance-independent and dependent traffic. *J. Lightwave Technology*, 17:48–57.

18. J.M. Simmons, A. Saleh (1999). The value of optical bypass in reducing router size in gigabit networks. In: Proceedings of ICC, 1:591–596.
19. R. Ul-Mustafa, A.E. Kamal (2006). Grooming of non-uniform traffic on unidirectional and bidirectional rings. *Computer Communications*, 29:1065–1078.
20. P.J. Wan, L. Liu, O. Frieder (1999). Grooming of arbitrary traffic in SONET/WDM rings. In: Proceedings of IEEE GLOBECOM, 1B:1012–1016.
21. P.J. Wan, G. Calinescu, L. Liu, O. Frieder (2000). Grooming of arbitrary traffic in SONET/WDM BLSRs. *IEEE J. Selected Areas in Communications*, 18: 1995–2003.
22. J. Wang, W. Cho, V. Vemuri, B. Mukherjee (2001). Improved approaches for cost-effective traffic grooming in WDM ring networks: ILP formulations and single-hop and multi-hop connections. *J. Lightwave Technology*, 19:1645–1653.
23. Y. Xu, S.C. Xu, B.X. Wu (2002). Traffic grooming in unidirectional WDM ring networks using genetic algorithms. *Computer Communications*, 25:1185–1194.
24. X. Zhang, C. Qiao (1998). An effective and comprehensive approach to traffic grooming and wavelength assignment in SONET/WDM rings. In: Proceedings of SPIE Conference On All-optical Networking, 3531:221–232.
25. X. Zhang, C. Qiao (2000). An effective and comprehensive approach for traffic grooming and wavelength assignment in SONET/WDM rings. *IEEE/ACM Transactions on Networking*, 8:608–617.
26. S. Zhang, B. Ramamurthy (2003). Dynamic traffic grooming algorithms for reconfigurable SONET over WDM networks. *IEEE J. on Selected Areas in Communications*, 21:1165–1172.

Traffic Grooming in Next-Generation SONET/SDH Networks

Smita Rai, Biswanath Mukherjee

8.1 Introduction

Telecommunication networks are generally segmented into a three-tier hierarchy: access, metropolitan, and long-haul (and further delineations are also possible). Long-haul/backbone networks span inter-regional/global distances (1000 km or more in the US context, and perhaps an order of magnitude shorter in many other densely populated regions) and are optimized for transmission and related costs. At the other end of the hierarchy are access networks, providing connectivity to a plethora of customers within close proximity. Straddled in the middle are metropolitan (metro) networks, averaging regions between 10–100 km and interconnecting access and long-haul networks [1].

Metro networks today are based on synchronous digital hierarchy (SDH)/synchronous optical network (SONET) ring architectures (for a detailed reference, please see [2]). Namely, smaller tributary rings, e.g., OC-3/STM-1 (155 Mbps) or OC-12/STM-4 (622 Mbps), aggregate traffic onto larger core inter-office (IOF) rings that interconnect central office (CO) locations at higher bit rates, e.g., OC-48/STM-16 (2.5 Gbps). Overall, SONET/SDH has been very successful in delivering the early wave of end-user connectivity, namely voice. However, with the stupendous growth in Internet data traffic, new metro solutions are required to offer superior price/performance alternatives for legacy SONET/SDH expansion. For example, new platforms must offer high bandwidth scalability and carry multiple protocols over a common infrastructure to reduce costs. Various “data-centric” solutions are emerging, driven by advances in high-density electronic integrated circuit (IC) technology, and “next-generation” SONET/SDH is the most prominent.

Next-generation SONET¹ provides enhancements over SONET functionality that are pertinent to traffic grooming. We briefly discuss these features

¹ We mostly use the term SONET in the rest of this chapter; however the ideas are equally applicable to SDH.

below and point out the benefit from a grooming point of view, before discussing grooming approaches. See also Chapter 4 and its references for discussion of these features.

NG-SONET is comprised of the following technologies:

Generic Framing Procedure (GFP) [3]: Support for GFP makes it possible to transport any protocol over SONET, including Gigabit-Ethernet, Fiber Channel, etc. GFP is a traffic-adaptation protocol for broadband transport applications. It provides a standard mapping of either a physical layer or logical link layer signal to a byte-synchronous channel such as SONET/SDH links or wavelength channels in an optical transport network (OTN). There are two methods for mapping protocols onto GFP: frame-mapped GFP and transparent-mapped GFP. Frame-mapped GFP, optimized for a packet-switched environment, is the transport mode used for Point-to-Point Protocol (PPP), IP, and Ethernet traffic. Transparent-mapped GFP, intended for delay-sensitive storage-area network (SAN) applications, is the transport mode used for Fiber Channel (FC), Fiber Connection (FICON), and Enterprise Systems Connection (ESCON) traffic. Figure 8.1 (from [4]) shows the mapping of GFP to SONET/SDH using VC.

Virtual Concatenation (VC/VCAT) [5]: Virtual concatenation lets carriers allocate bandwidth within a SONET pipe to different services. For instance, the bandwidth can be allocated in 51.84 Mbps chunks (higher-order VCAT) based on the number of STS-1 channels used.

Link-Capacity Adjustment Scheme (LCAS) [6]: LCAS has been developed to allow the dynamic allocation of bandwidth to different services. It allows changes in bandwidth in response to customers requesting increases or decreases of bandwidth. LCAS, built on virtual concatenation, is a two-way signaling protocol that runs continuously between the source and destination of a bandwidth pipe. LCAS allows network operators to adjust the pipe capacity while it is in use (on the fly). It increases the possibility for on-demand traffic provisioning and on-line traffic grooming/re-grooming and makes SONET/SDH-based optical WDM network more data friendly.

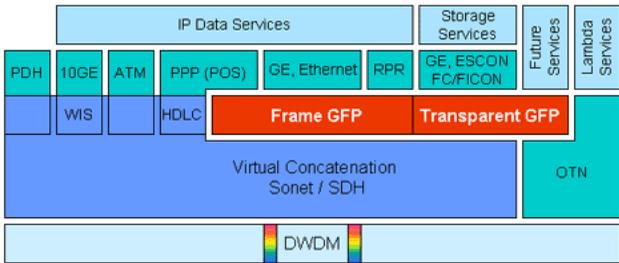


Fig. 8.1. GFP mapping to SONET/SDH.

VCAT, a feature that allows “inverse multiplexing” of traffic flows over multiple paths, has important consequences for traffic grooming, and we give a brief overview of this technology in the following section.

8.2 Virtual Concatenation

As the amount of data traffic rapidly increases, the inefficiency of transporting packet data through a SONET/SDH frame emerges as a major concern when network operators try to optimize the usage of their current bandwidth to support various types of new services and applications, based on IP, Frame Relay, Ethernet, Fiber Channel, etc.

In the traditional SONET/SDH multiplexing hierarchy, the frames of multiple low-speed traffic streams (say, STS-1 frame, approx. 51.84 Mbps) are combined to form the frame of a high-speed traffic stream. In order to support high-speed traffic from the same client source, e.g., a broadband ATM switch, N “contiguous” lower-order SONET/SDH payload containers are merged into one of greater capacity. This is called SONET/SDHs concatenation technique.

Usually, SONET/SDH’s concatenation is implemented at certain speeds, such as STS-3c, STS-12c, STS-48c, etc., which leads to a tiered bandwidth-allocation mechanism for different client services. Unfortunately, although “contiguous” and “tiered” concatenation is simple for implementation, it is not very flexible and not very efficient, especially in a multi-service network environment. From a single network node perspective, traffic streams from different client network equipment are to be discretely mapped into different tiers of SONET/SDH bandwidth trunks (data containers), which may result in large capacity wastage. For example, carrying a Gigabit Ethernet connection using a concatenated OC-48 pipe (approx. 2.5 Gbps) will lead to 60% bandwidth wastage.

From a network perspective, the time-slot contiguous requirement of a SONET/SDH concatenated channel imposes a constraint for traffic provisioning and may degrade network performance in a dynamic traffic environment where network resources are easy to be fragmented. The constraint also makes it more difficult for a network operator to perform efficient traffic grooming, i.e., packing different low-speed traffic streams onto high-capacity wavelength channels.

Virtual concatenation (VC or VCAT) will help a SONET/SDH-based optical network to carry data traffic in a finer granularity and hence use link capacity more efficiently. The basic principle of virtual concatenation [5] is quite simple. A number of smaller containers, which are not necessarily contiguous, are concatenated and assembled to create a bigger container that carries more data per second. Depending on a network’s switching granularity, virtual concatenation is possible for small container size from VT-1.5 up to STS-3c.

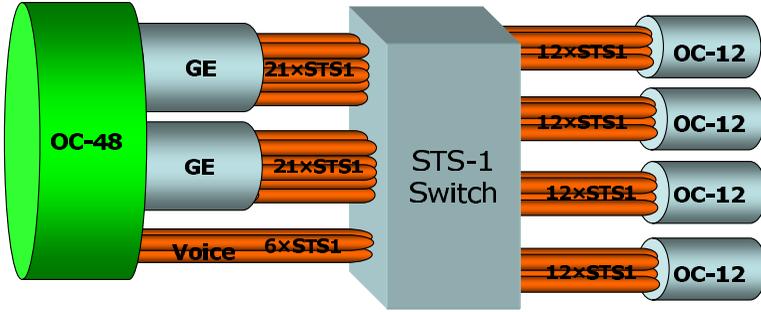


Fig. 8.2. An example of using VC to support different network services

Figure 8.2 (from [7]) shows an example of how to support multiple services using a single OC-48 channel through virtual concatenation. In Fig. 8.2, an OC-48 channel is used to carry two Gigabit Ethernet traffic streams and six STS-1 TDM voice traffic streams. Through a STS-1 switch, the traffic can be switched onto different OC-12 pipes, and these OC-12 pipes can be sent through the network over various routes. Figure 8.2 also illustrates the potential load-balancing benefit that VC can provide to a transport network.

One important thing to note is that, when multiple traffic streams from one client are sent over different routes, the VC mappers at the destination nodes need to compensate for the differential delay between the bifurcated streams when they are reconstructed at the destination node. Currently, such a typical, commercially available device may support up to 50 ms (+/-25 ms) delay with external RAM, which is equivalent to a 10,000 km difference in route length [8]. In general, a virtually concatenated SONET/SDH channel made up of $N \times$ STS-1 is transported as individual STS-1s across the network; and, at the receiver, the individual STS-1s are re-aligned and sorted to recreate the original payload. Figure 8.3 (from [1]) shows an overview of a multi-service SONET/SDH-based optical network employing VC technology. VC can be supported at the edge optical crossconnects (OXC) (in port cards) or in separate traffic-aggregation network elements connecting client network equipment and OXC.

8.2.1 Benefits of Virtual Concatenation

From a network perspective, the SONET/SDH VC technique can provide the following benefits to an optical network:

1. *Relax time-slot alignment and continuity constraints.* Instead of aligning to particular time-slots and consisting of N contiguous STS-1 time-slots within a wavelength, a high-speed STS-N channel could be constructed from any N STS-1 time-slots and carried by different wavelength channels.

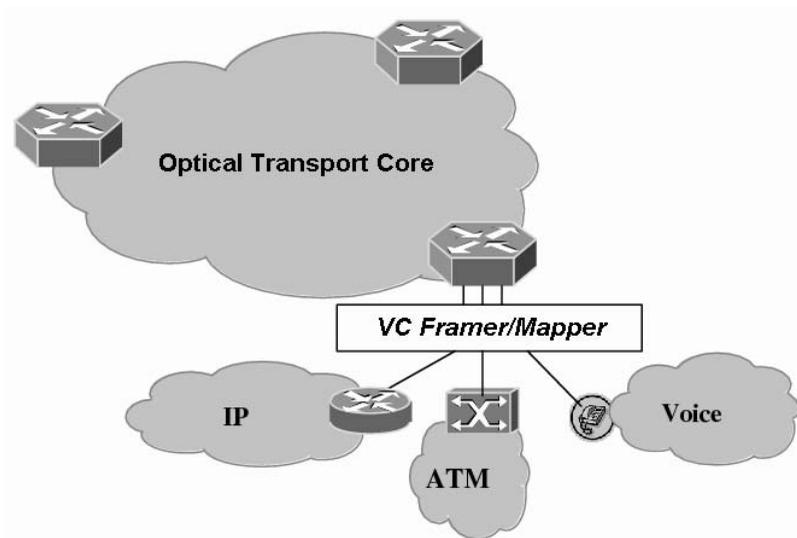


Fig. 8.3. An overview of a VC-enabled multi-service optical network

2. *More efficiently utilize channel capacity to support multiple types of data and voice services.* Instead of mapping data traffic (packet/cell/frame) into SONET/SDH frames in a discrete tiered manner, optical networks can carry data traffic in a more resource-efficient way. Traffic granularity can be increased in the unit of 1.6 Mbps (VT-1.5) in a metro-area network, and 48 Mbps (STS-1) or 150 Mbps (STS-3c) in an optical backbone network.
3. *Bifurcate traffic streams to balance network load.* With VC, it is possible to split a high-speed traffic stream into multiple low-speed streams and route them separately through the network. This enables the traffic to be distributed across the network more evenly; and, hence, it can improve the network's blocking performance.

8.3 Algorithm for Provisioning with Virtual Concatenation

The benefits of SONET/SDH VC to an optical WDM mesh network (shown in Fig. 8.4) under dynamic traffic environment is investigated via simulations in [9]. Connections with different bandwidth granularities are assumed to come and leave the network, one at a time, following a Poisson arrival process and negative-exponential-distribution holding time.

To distinguish the benefit effects, two types of traffic pattern are studied (pattern I and II), one consisting of five service classes and the other having ten service classes. Capacity of each wavelength channel is assumed to be OC-192. In traffic pattern I, data rates for each service class are approximately 51

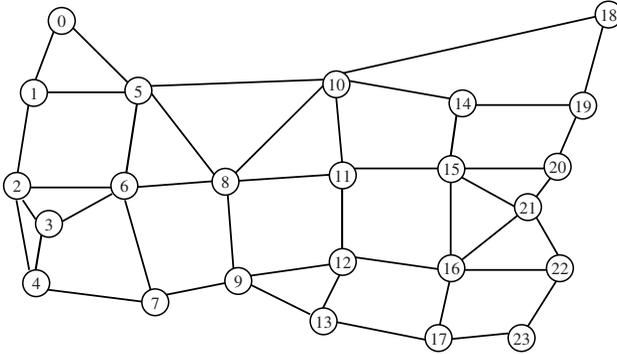


Fig. 8.4. A 24-node example network topology

Mbps, 153 Mbps, 622 Mbps, 2.5 Gbps, and 10 Gbps, which can be perfectly mapped into the tiered SONET/SDH containers and their corresponding optical carriers, i.e., OC-1, OC-3c, OC-12c, OC-48c, and OC-192c. The service characteristic of pattern II, i.e., service classes, service rate, and corresponding SONET/SDH containers with or without VC is shown in Table 8.1.

When traffic bifurcation is needed, a simple route-computation and traffic-bifurcation heuristic is applied to a connection request. The problem of finding the set of minimal-cost routes with enough aggregated capacity for a request between a given node pair in a network is NP-Complete. The heuristic works as follows:

- Step 1: A shortest path is computed according to network administrative cost between the node pair.
- Step 2: The bandwidth of the route is calculated. The bandwidth of the route is constrained by the link along the route, which has minimal free capacity. Then, update the available capacity of the links along the route (i.e., decrease the available capacity by the minimal capacity).
- Step 3: Remove the link without free capacity and repeat Steps 1 and 2 until the connection can be carried by the set of routes computed or no more routes exist for the connection.

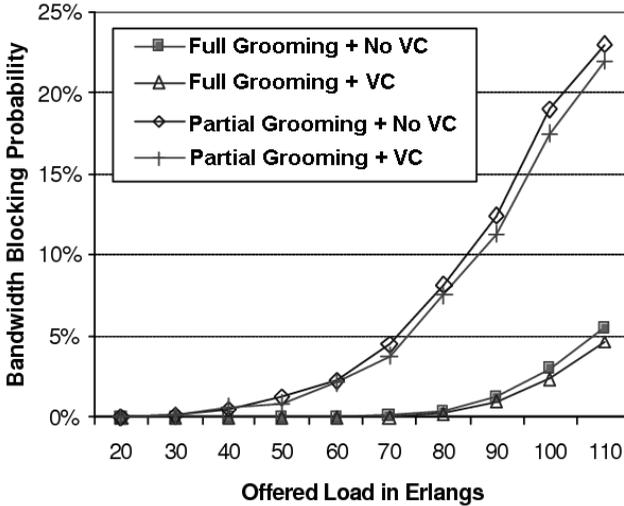
Note that, depending on the implementation, the VC mappers at the receiver end nodes may only be able to handle a limited number of routes

Table 8.1. Traffic pattern II: ten service classes

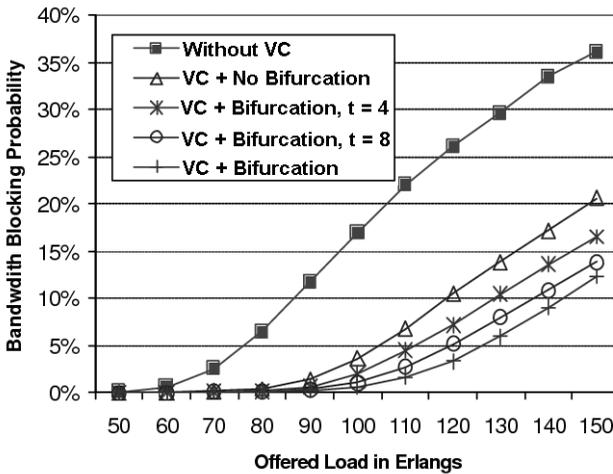
Class	Rate	Without VC	With VC	Class	Rate	Without VC	With VC
1	50 M	STS-1	STS-1	6	600 M	STS-12c	STS-12
2	100 M	STS-3c	STS-2	7	1 G	STS-48c	STS-21
3	150 M	STS-3c	STS-3	8	2.5 G	STS-48c	STS-48
4	200 M	STS-12c	STS-4	9	5 G	STS-192c	STS-96
5	400 M	STS-12c	STS-8	10	10 G	STS-192c	STS-192

(denoted by t) for a single connection, in which case, the connection will be blocked after t routes have been examined.

Figure 8.5 illustrates network performance in terms of bandwidth blocking probability (BBP) as a function of network offered load in Erlangs. BBP is considered as the measurement metric because connections from different service classes may have different bandwidth requirements. Network offered load



(a) Five service classes.



(b) Ten service classes.

Fig. 8.5. Illustrative numerical results

is normalized to the unit of OC-192. A 24-node, 43-bidirectional-link network topology is used in our study (see Fig. 8.4). Each link has 16 wavelengths.

Figure 8.5(a) shows the network performance for traffic pattern I with or without employing VC. Two types of network configurations are examined, i.e., all nodes are either equipped with STS-1 full-grooming switches or partial-grooming switches. Note that, in a partial-grooming switch, only a limited number of wavelength channels (6 in our simulation) can be switched to a separate grooming switch (or grooming fabric within an OXC) to perform traffic grooming.

As one can observe from Fig. 8.5(a), there is around 5–10 percent network performance gain through VC technique. In traffic pattern I, every service class can be perfectly mapped into one of the tiered SONET/SDH containers, and we assumed that no traffic bifurcation is allowed. Therefore, the performance improvement shown in Fig. 8.5(a) comes solely from the capability of eliminating the time-slot alignment and contiguity constraints provided by VC to the network.

Figure 8.5(b) shows how VC can significantly improve network performance when the network needs to support data-oriented services with different bandwidth requirements under the network configuration in which full-grooming OXCs are used at every node. It is straightforward to see that BBP is significantly reduced by employing VC. Meanwhile, more improvement can be achieved by allowing a simple traffic-bifurcation scheme. Note that, in the study, a connection will be bifurcated if and only if no single route with enough capacity exists. The results from different values of t have been examined and some of them are shown in Fig. 8.5(b), i.e., 4, 8, and unlimited. It can be expected that more advanced network load-balancing and traffic-bifurcation approaches can further optimize network throughput.

The benefits quantitatively demonstrated above arise from more efficient grooming that exploits the flexibility provided by Virtual Concatenation (VC). Since a connection can be split up, and routed across multiple paths, and there are fewer constraints on time-slot alignment, grooming techniques that pack connections more effectively can be developed. SONET/SDH VC could also benefit an optical transport network on other aspects, such as network compatibility, network resiliency, network management and control, etc. VC works across legacy networks. Only the end nodes of the network are aware of the containers being virtually concatenated. In terms of resiliency, since individual members of a virtually concatenated channel may be carried through different routes, a network failure may only affect partial bandwidth of a connection service. In this case, an unprotected, best-effort connection may still get the service under the reduced bandwidth before the network failure is fixed, and a high-priority connection can still be provided partial service before the network protection/restoration operation is active to restore the full service. Reference [10] explores reliability-based provisioning in next-generation SONET/SDH networks addressing some of these concepts of degraded service and expected bandwidth for a multi-path connection.

Reference [11] investigates novel schemes for protecting multi-path connections by introducing concepts such as “intra-connection” sharing apart from the conventional “inter-connection” sharing.

8.4 Next-Generation SONET in Interconnected Rings

In this section, we explore the problem of provisioning connections in dual-node-interconnected SONET/SDH ring networks. Dual-node interconnection using drop-and-continue architecture (for more details, please refer to [12, 13]) is the de-facto SONET standard for connecting rings – Bi-directional Line Switched Rings (BLSR) or Uni-directional Path Switched Rings (UPSR) – to ensure survivability against all single-link and single-node failures. Similar recommendations exist for SDH networks [13]. This architecture, combined with legacy SONET time-slot assignment rules for contiguous-concatenated signals (equivalent rules exist for SDH), places some unique constraints on provisioning connections dynamically on such ring networks. We explore the benefits of employing virtual concatenation and grooming high-bandwidth connections over multiple paths in such networks.

We developed an algorithm called `PATH.FINDER` [14] to find valid routes in such a network of interconnected rings. The basic intuition behind the algorithm is as follows. Within a ring, there are not too many choices of alternate paths. For an intra-ring demand, we can only choose one of two directions in which the connection should travel for BLSR rings. However, when an inter-ring connection has to travel through several intermediate rings, the choice of the next-hop ring becomes crucial, at each stage. We capture this intuition by constructing a simplified graph from the current state of the network.

This graph contains the source and destination of the connection request, as well as all the interconnection pairs between each ring. These dual-node interconnection pairs on a ring are collapsed into one node in the simplified graph (see Fig. 8.6). We add an edge between two nodes in the simplified graph, called an auxiliary graph, if there exists a valid path between them in the original network (see Fig. 8.7).

Note that the above construction is not necessary for demands whose source and destination nodes are within the same ring. For BLSR rings, the direction in which to send the demand can be chosen based on some path metric.

We also developed an intelligent cost and time-slot assignment algorithm (called `SmartCost`) to minimize fragmentation while provisioning a connection. For details, the reader is referred to [14].

To study the properties of our provisioning algorithm, we simulated a dynamic network environment. The network topology is shown in Fig. 8.8. It is composed of 8 OC-192 BLSR rings with a total of 50 nodes and 12 interconnection pairs. The large number of interconnection pairs reduces the effect of the drop-and-continue penalty, by allowing traffic to be load balanced

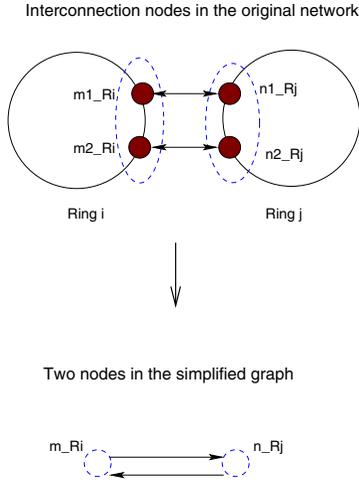


Fig. 8.6. Graph transformation

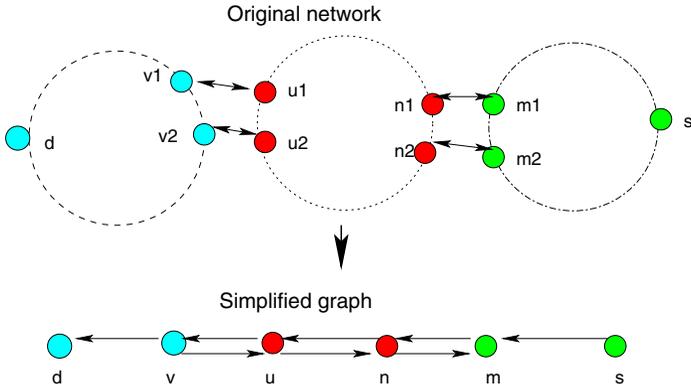


Fig. 8.7. Auxiliary graph

over different interconnection pairs, as was suggested in [15]. The connection-arrival process is Poisson and the connection-holding time follows a negative exponential distribution, in which source and destination nodes are selected uniformly from among the nodes in the network with inter-ring demands and intra-ring demands being in the ratio 25:75. The traffic mix is as follows: $STS-1:STS-3c:STS-12c:STS-48c = 100:50:10:1$. This is based on the assumption that lower-bandwidth requests come in at a larger proportion than higher-bandwidth requests. All connections are bi-directional. Note that the contiguous concatenation and alignment constraints are too restrictive for STS-48c in the network topology using OC-192 rings. Such a request can

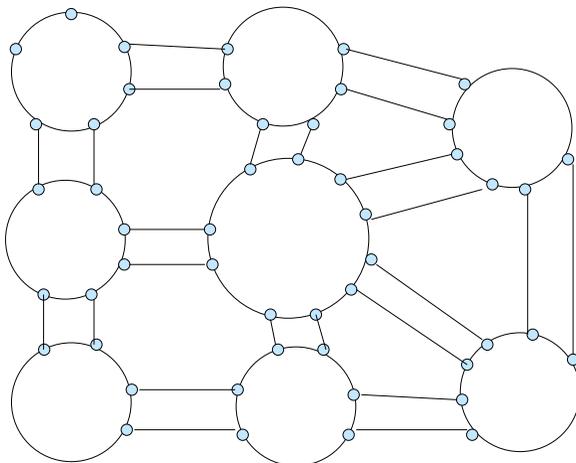


Fig. 8.8. Ring topology

use slots starting at only two time slots, namely 1 and 49 on any link, and all the adjacent 47 slots must also be free.

We compared our SmartCost algorithm to a shortest-hop approach with first-fit to provision paths. In first-fit time-slot assignment, the connection is assigned to the first free set of time slots which satisfy the contiguity and alignment requirements. Both algorithms use our graph abstraction (our PATH_FINDER module) to select a route out of a set of valid routes; the differences lie in how cost is assigned to a path and how time slots are assigned.

As a measure of efficient capacity utilization, we compare the blocking ratio (also called blocking probability) which is the ratio of the number of connections blocked for a particular granularity to the number of requests for that granularity, as well as the overall bandwidth blocking ratio. For all our simulation results, we find the 95% confidence intervals according to the t -distribution [16] by performing several trials of each experiment.

We observed a high blocking ratio of STS-48 with both the algorithms (Fig. 8.9), true to our intuition that contiguous concatenation imposes significant penalty on higher-bandwidth requests. We explored how next-generation SONET technologies, such as VCAT, can help to significantly improve the blocking ratio of these high-bandwidth requests by grooming them over multiple paths. From our simulation results in Fig. 8.9, we observed that a significant chunk of high-bandwidth connections, such as STS-48, are blocked without the facility of VCAT. We incorporated virtual concatenation in our algorithms, allowing an STS-48 signal to be broken up into 48 STS-1 signals, and routed independently in the network. The other connections are carried using contiguous concatenation as before. We observe substantial improvement in the blocking ratio of STS-48 requests (Fig. 8.9), and significant improvement in capacity utilization. The total bandwidth blocking ratio

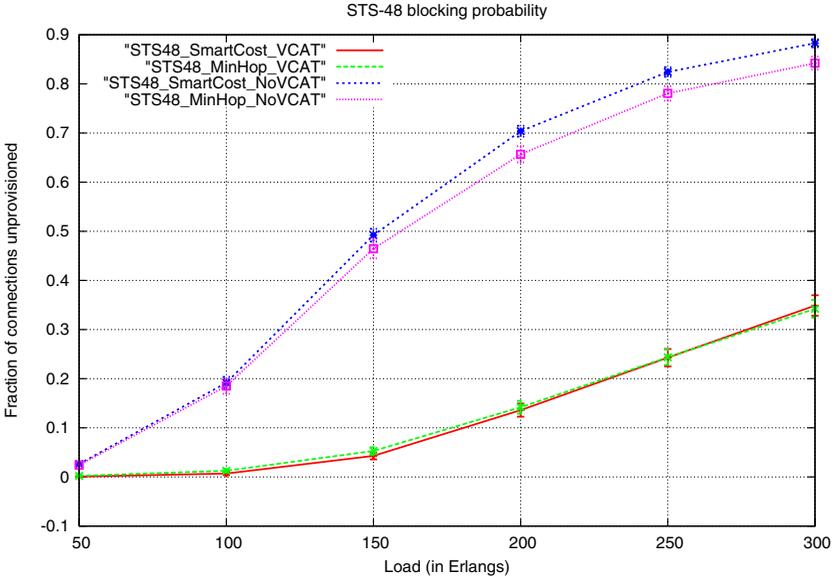


Fig. 8.9. STS-48 blocking ratio with and without VCAT

with the introduction of VCAT decreases to 1/3 of the value without it at intermediate loads (Fig. 8.10). There is a slight increase in the fraction of STS-12c connections being provisioned (Fig. 8.11), since capacity is used up by STS-48 bandwidth connections; however the overall bandwidth blocking

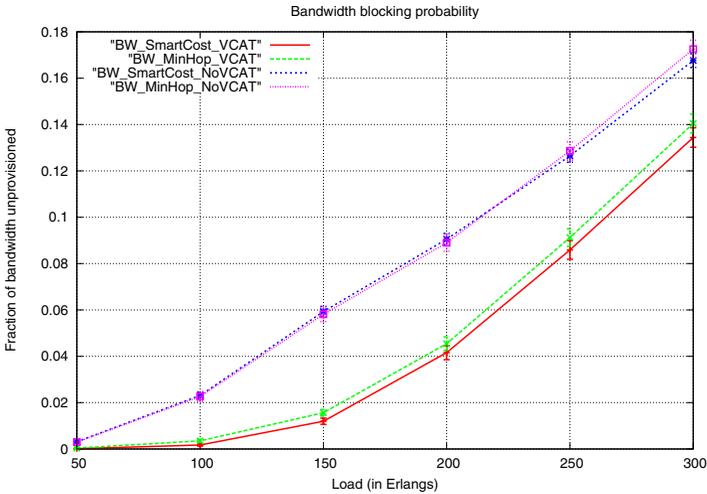


Fig. 8.10. Bandwidth blocking ratio with and without VCAT

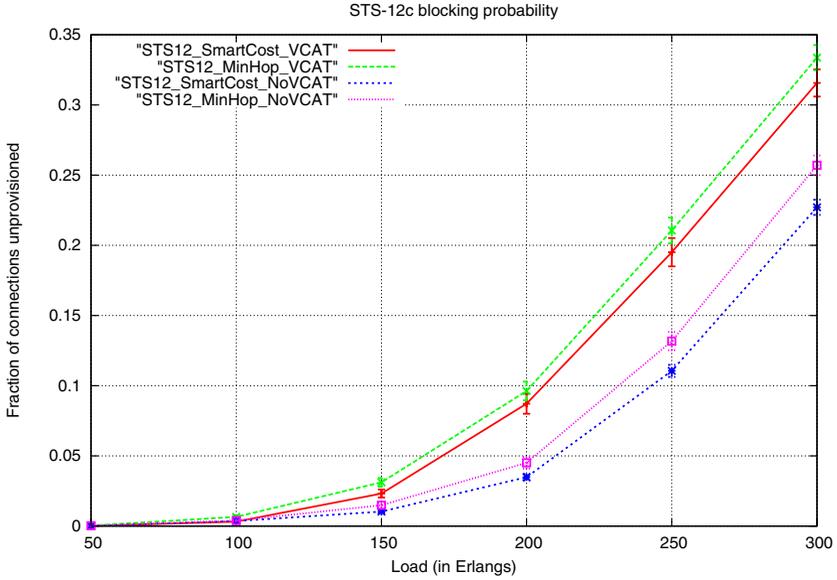


Fig. 8.11. STS-12c blocking ratio with and without VCAT

ratio improves significantly. The blocking ratio of lower-bandwidth requests shows negligible difference (Fig. 8.12), and for STS-1 the ratio is effectively zero for all loads and all strategies. SmartCost algorithm performs better than the minimum-hop scheme; however, substantial improvement due to flexible

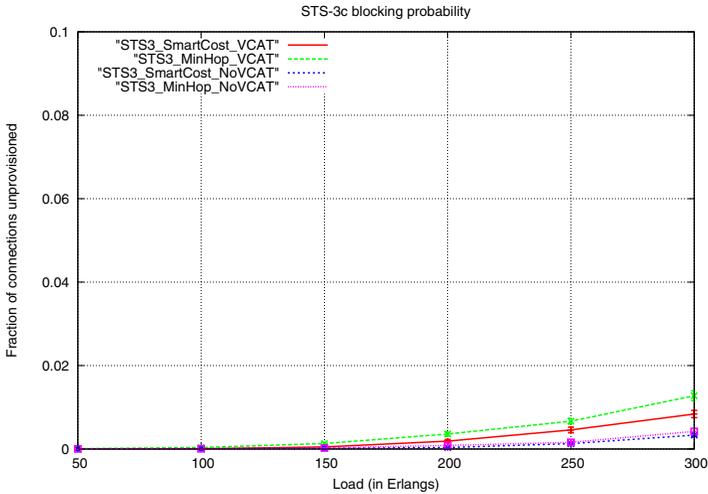


Fig. 8.12. STS-3c blocking ratio with and without VCAT

concatenation is observed when comparing any approach without VCAT to a VCAT-enabled approach.

With the ability to split up and route a connection over different paths, there arises the problem of differential delay among the different paths. The standards allow a delay up to 256 ms, but practical values range from 60 to 100 ms [17]. Addressing differential delay issues are out of scope of this chapter; the interested reader is referred to [18, 19] for some work on differential-delay-based provisioning. Note that, since we focus on a network of rings, if the STS-48c demands are split up using VCAT, for an intra-ring demand, it is expected that the differential delay would not pose a serious problem (there are only two directions to choose from and a BLSR ring can have a maximum 16 nodes, with 1200 km of total fiber [2]). For an inter-ring demand, in our algorithms, we do not allow paths to loop around the same ring multiple times. For instance, paths such as $Ring_A \rightarrow Ring_B \rightarrow Ring_C \rightarrow Ring_B$ are not permitted. Due to the limited number of rings, and limits on the maximum number of nodes on a ring as well as the total fiber, we expect that differential delay issues would be less problematic. Further exploration of delay issues is an open problem for future research.

8.5 Conclusion

Virtual concatenation (VC / VCAT) and link-capacity adjustment scheme (LCAS) can help a SONET/SDH-based optical network to evolve towards a data-centric intelligent automatically switched optical network. They offer the following benefits: (a) relaxing the time-slot continuity and alignment constraints of traditional SONET/SDH concatenation; (b) improving bandwidth efficiency of a wavelength channel; (c) enabling inverse-multiplexing (i.e., traffic bifurcation) and load balancing; and (d) improving service resilience. By loosening the constraints on time-slot assignment and by allowing more flexibility in provisioning, these benefits allow for the development of efficient traffic grooming techniques that can pack connections more effectively in SONET/SDH pipes. We studied these benefits in a general mesh network as well as in a network of interconnected rings using dual-node interconnection. We presented illustrative numerical results to demonstrate the significant benefits which can be obtained by employing virtual concatenation.

Our simple provisioning algorithms that groom connections onto multiple paths, exploiting the inverse multiplexing capability of VCAT, show significant performance gains, when compared to approaches using legacy SONET/SDH.

Moreover, built on virtual concatenation, LCAS allows a network operator to adjust the pipe capacity while it is in use (on the fly). This increases the possibility for on-demand traffic provisioning and it makes SONET/SDH-based optical WDM networks more data friendly.

References

1. B. Mukherjee, *Optical WDM Networks*, Springer, Optical Network Series, 2006.
2. W. Goralski, *SONET/SDH*, 3rd edition, McGraw Hill, 2002.
3. ITU-T, Recommendation G.704/Y.1303, "Generic Framing Procedure (GFP)," 2001.
4. S. Stanley, "Making SONET Ethernet-friendly," Lightreading report, March 2003.
5. ITU-T, Recommendation G. 707, "Network node interface for the synchronous digital hierarchy (SDH)," April 2002.
6. ITU-T, Recommendation G.7042/Y.1305, "Link capacity adjustment scheme (LCAS) for virtual concatenated signals," 2001.
7. S. Stanley, "Next-gen SONET silicon," Lightreading report, June 2002.
8. T. Hills, "Next-gen SONET," Lightreading report, May 2002.
9. K. Zhu, J. Zhang, and B Mukherjee, "Ethernet-over-SONET (EoS) over WDM in Optical Wide-Area Networks (WANs): Benefits and Challenges," *Photonics Network Communications*, vol. 10, no. 1, pp. 107–118, July 2005.
10. S. Rai, O. Deshpande, C. Ou, and B. Mukherjee, "Reliable multi-path provisioning for high-capacity optical backbone mesh networks," *Proc., IEEE International Conference on Communications*, June 2005. Extended version to appear in *IEEE/ACM Transactions on Networking*, vol. 15, no. 4, August 2007.
11. C. Ou, L. H. Sahasrabudde, K. Zhu, C. U. Martel, and B. Mukherjee, "Survivable virtual concatenation for data over SONET/SDH in optical transport networks," *IEEE/ACM Transactions on Networking*, vol. 14, no. 1, Feb. 2006.
12. G. W. Ester, "SONET: Interconnecting rings using drop and continue," *Telephone Engineer and Management*, June 1, 1993.
13. ITU-T Recommendation, "Interworking of SDH network protection architectures," G.842, April 1997.
14. S. Rai, L. Song, C. -F. Su, T. Hamada, and B. Mukherjee, "A novel provisioning framework for dual-node interconnected SONET/SDH rings," *Proc., Optical Fiber Communications*, March 2006.
15. B. T. Doshi, S. Dravida, P. Harshavardhana, P. K. Johri, and R. Nagarajan, "Dual (SONET) ring interworking: High penalty cases and how to avoid them," *Proc., International Teletraffic Congress ITC 15*, Elsevier 1997.
16. K. S. Trivedi, *Probability and Statistics with Reliability, Queuing, and Computer Science Applications*, Englewood Cliffs, NJ: Prentice-Hall, 1982.
17. G. Bernstein, B. Rajagopala, and D. Saha, *Optical Network Control: Architecture, Protocols and Standards*, Addison-Wesley, July 2003.
18. A. Srivastava, S. Acharya, M. Alicherry, B. Gupta, and P. Risbood, "Differential Delay aware Routing for Ethernet over SONET/SDH," *Proc., INFOCOM'05*, March 2005.
19. S. Huang, S. Rai, and B. Mukherjee, "Survivable differential delay aware multi-service over SONET/SDH networks with virtual concatenation," *Proc., Optical Fiber Communications (OFC) Conference*, March 2007.

Mathematical Programming Approaches

JianQiang Hu

9.1 Introduction

There are two basic architectures used in WDM networks: ring and mesh. The majority of optical networks in operation today have been built based on the ring architecture. However, carriers have increasingly considered the mesh architecture as an alternative for building their next generation networks. Various studies have shown that mesh networks have a compelling cost advantage over ring networks. Mesh networks are more resilient to various network failures and also more flexible in accommodating changes in traffic demands (e.g., see [4, 6, 17] and references therein). In order to capitalize on these advantages, effective design methodologies are required.

The grooming problem in an optical mesh network in fact involves traffic grooming, routing, and wavelength assignment. The problem of traffic grooming and routing (GR) for mesh networks is to determine how to efficiently route traffic demands and at the same time to combine lower-rate (sub-wavelength) traffic demands onto a single wavelength. On the other hand, the problem of wavelength assignment (WA) is to determine how to assign specific wavelengths to lightpaths, usually under the wavelength continuity constraint. In previous studies on the routing and wavelength assignment (RWA) problem for mesh networks (e.g., see [12, Chapter 8] and references therein), the issue of traffic grooming has largely been ignored, i.e., it has been assumed that each traffic demand takes up an entire wavelength. In practice, this is hardly the case, and networks are typically required to carry a large number of lower-rate traffic demands. We should point out that the grooming problem is an NP-complete problem; therefore, heuristics and approximations are usually needed for solving this problem.

The traffic grooming problem for mesh networks has only been considered recently (e.g., see [5, 7, 8, 10, 16]). The objective considered in [16] is either to maximize the network throughput or to minimize the connection-blocking probability, which are operational network-design problems. A different operational network-design problem is considered in [10], in which a Lagrangian

relaxation based method is proposed to minimize the total cost associated with links. Alternatively, a strategic network-design problem is to minimize the total network cost. Typically, the cost of a nation-wide optical network is dominated by optical transponders and optical amplifiers. If one assumes that the fiber routes are fixed (e.g., see [14, 15] for some recent work on the problem of amplifier allocation), then the amplifier cost is constant, in which case one should concentrate on minimizing the number of transponders in the network. Ding and Hamdi [5] proposed a heuristic algorithm to minimize the number of transponders as well as the number of wavelengths in mesh networks. In [8], Lee and Park proposed a genetic algorithm to minimize a combination of the number of transponders and the number of wavelengths. In [7], Hu and Leida first formulated the grooming problem as an integer linear programming (ILP) problem, then decomposed it into two smaller problems: the GR problem and the WA problem.

This chapter is mainly based on the work of [7], where the problem of traffic grooming, routing, and wavelength assignment (GRWA) was considered with the objective of minimizing the number of transponders in the network. We first formulate the GRWA problem as an integer linear programming (ILP) problem. Unfortunately, the resulting ILP problem is usually very hard to solve computationally, in particular for large networks. To overcome this difficulty, we then propose a decomposition method that divides the GRWA problem into two smaller problems: the traffic grooming and routing (GR) problem and the wavelength assignment (WA) problem. In the GR problem, we only consider how to groom and route traffic demands onto lightpaths (with the same objective of minimizing the number of transponders) and ignore the issue of how to assign specific wavelengths to lightpaths. Similar to the GRWA problem, we can formulate the GR problem as an ILP problem. The size of the GR ILP problem is much smaller than its corresponding GRWA ILP problem. Furthermore, we can significantly improve the computational efficiency for the GR ILP problem by relaxing some of its integer constraints, which usually leads to quite good approximate solutions for the GR problem. Once we solve the GR problem, we can then consider the WA problem, in which our goal is to derive a feasible wavelength assignment solution.

We note that the WA problem has been studied by several researchers before (e.g., see [1, 2, 3, 9, 11, 12, 13] and references therein). However, the objective in all these studies has been to minimize the number of wavelengths required in a network, in some cases by using wavelength converters. In general, the use of additional wavelengths in a network only marginally increases the overall network cost as long as the total number of wavelengths used in the network does not exceed a given threshold (the wavelength capacity of a WDM system). This is mainly because the amplification cost is independent of the number of wavelengths. In recent years, the wavelength capacity for optical networks has increased dramatically. For example, with most advanced techniques, a single WDM system on a pair of fibers can carry up to 160

10G-wavelengths or 80 40G-wavelengths. Of course, once the wavelength capacity is exceeded, then a second parallel system (with another set of optical amplifiers) needs to be built, which would then substantially increase the network cost. Therefore, assuming a single WDM system on all fiber routes fixes the amplifier cost, then one should focus on minimizing the number of transponders in the network, which is already taken into consideration in the GR problem. In this setting, the objective in our WA problem is to find a feasible wavelength assignment solution under the wavelength capacity constraint.

It is clear that, in general, the decomposition method would not yield the optimal solution for the GRWA problem. However, we will provide a sufficient condition under which we show that the decomposition method does produce an optimal solution for the GRWA problem. This is achieved by developing a simple algorithm that, under this sufficient condition, finds an optimal wavelength assignment.

The rest of this chapter is organized as follows. In Section 9.2, we present the GRWA problem and demonstrate how it can be formulated as an ILP problem. In Section 9.3, we first present our decomposition method. We then provide an ILP formulation for the GR problem and develop an algorithm for solving the WA problem. We also discuss under what condition the decomposition method produces an optimal solution for the GRWA problem. Some numerical results are provided in Section 9.4.

9.2 The GRWA Problem

An optical mesh network architecturally has two layers: a physical layer and an optical layer. The physical layer consists of fiber spans and nodes and the optical layer consists of lightpaths (optical links) and a subset of nodes contained in the physical layer. A lightpath in the optical layer is a path connecting a pair of nodes via a set of fiber spans in the physical layer. Throughout this chapter, we assume that lightpaths and their routes in the physical layer are given. In practice, the selection of lightpaths is another important design issue that needs to be addressed, which is beyond the scope of this chapter.

We use graph $G_f = (V_f, E)$ to represent the physical layer, where E is the set of edges representing fiber spans and V_f is the set of nodes representing locations which are connected via fiber spans. We use graph $G_o = (V_o, L)$ to represent the optical layer, where L is the set of edges representing lightpaths and $V_o \subset V_f$ is a subset of locations that are connected via lightpaths. Each edge in L corresponds to a path in G_f . Here we treat each lightpath as a logical connection between a pair of nodes (not just a single wavelength); therefore, one lightpath can contain multiple wavelengths. For ease of exposition, we first assume that G_o is a directed graph (i.e., the lightpaths are unidirectional). The extension to the undirected graph case is quite straightforward

and will be discussed later in this section (basically, we can simply replace every undirected edge with two directed edges).

The GRWA problem studied in this chapter can be described as follows. Assuming that a set of traffic demands are given (some of them are of low rate, i.e., sub-wavelength), our goal is to find an optimal way to route and groom these demands in the optical layer, G_o , and also to assign a set of specific wavelengths to each lightpath so that the total number of transponders required is minimized. There are two key constraints we need to take into consideration in this problem: (1) the wavelength capacity constraint for each fiber span, and (2) the wavelength continuity constraint for every lightpath, i.e., the same wavelength(s) needs to be assigned to a lightpath over the fiber spans it traverses (we assume that wavelength converters are not used; therefore the wavelength continuity constraint is required). In this problem setting, the number of transponders required for each lightpath is equal to twice the number of wavelengths assigned to it (one transponder for each end of each wavelength on a lightpath). Therefore, by grooming several low rate demands onto a single wavelength, we can potentially reduce the total number of wavelengths required by the lightpaths, thus the number of transponders.

The GRWA problem can be formulated as an integer linear programming (ILP) problem. First, we need to introduce some necessary notation:

W : the set of wavelengths available on each fiber;

D : the set of traffic demands;

g : the capacity of a single wavelength;

s_d : the size of demand $d \in D$;

A : $= [a_{v,l}]_{|V_o| \times |L|}$, the node-edge incidence matrix of graph G_o , where $a_{v,l} = 1$ if lightpath l originates from node v , -1 if lightpath l terminates at node v , and 0 otherwise;

B : $= [b_{e,l}]_{|E| \times |L|}$, the fiber-lightpath incidence matrix, where $b_{e,l} = 1$ if fiber span e is on lightpath l , and 0 otherwise;

u_d : $= [u_{v,d}]_{v \in V_o}$, the source-destination column vector for $d \in D$, where $u_{v,d} = 1$ if v is the starting node of d , -1 if v is the end node of d , and 0 otherwise;

x_d : $= [x_{l,d}]_{l \in L}$, the column vector containing lightpath routing variables for $d \in D$, where $x_{l,d} = 1$ if demand d traverses lightpath l , and 0 otherwise;

y_w : $= [y_{l,w}]_{l \in L}$, the column vector containing wavelength assignment variables for $w \in W$, where $y_{l,w} = 1$ if wavelength w is assigned to lightpath l , and 0 otherwise (note that in our setting each lightpath l is treated as a logical connection between a pair of nodes, hence it can be assigned with multiple wavelengths, i.e., it is possible that $\sum_{w \in W} y_{l,w} \geq 1$);

$\mathbf{1}$: $= [1, 1, \dots, 1]$, the unit column vector of appropriate size.

Then the GRWA problem can be formulated as the following ILP problem (which we shall refer to as the GRWA ILP problem):

$$\min \sum_{w \in W, l \in L} y_{l,w}$$

$$\text{s.t. } Ax_d = u_d \quad d \in D \quad (9.1)$$

$$By_w \leq \mathbf{1} \quad w \in W \quad (9.2)$$

$$\sum_{d \in D} s_d x_{l,d} \leq g \sum_{w \in W} y_{l,w} \quad l \in L \quad (9.3)$$

x and y are binary variables.

where the objective function $\sum_{w \in W, l \in L} y_{l,w}$ is the total number of wavelengths assigned to all lightpaths, which is equivalent to minimizing the total number of transponders needed. The three constraints are:

- Equation (9.1) is the flow balance equation, which guarantees that the lightpaths selected based on x_d constitute a path from the starting node of d to the end node of d .
- Equation (9.2) implies a single wavelength along each fiber span can be assigned to no more than one lightpath.
- Equation (9.3) is the capacity constraint for lightpath l , since $\sum_{d \in D} s_d x_{l,d}$ is the total amount of demands carried by lightpath l and $g \sum_{w \in W} y_{l,w}$ is the total capacity of lightpath l .

We refer to the type of the network considered above as the basic model. There are several variations of the basic model, which include:

1. networks with both protected and unprotected demands;
2. networks in which lightpaths are undirected;
3. networks with non-homogeneous fibers where different types of fiber may have different wavelength capacities; and
4. networks in which demand exceeds a single WDM system per fiber pair.

9.3 A Decomposition Method

In the previous section, we formulated the GRWA problem as an ILP problem; however, it may not be computationally feasible to solve the ILP problem, particularly for large networks (e.g., see numerical results in Section 9.4). Therefore, it is necessary to find more efficient ways to solve the GRWA problem. In this section, we propose a decomposition method that divides the GRWA problem into two smaller problems: the traffic grooming and routing (GR) problem and the wavelength assignment (WA) problem. In the GR problem, we only consider how to groom and route demands over lightpaths and ignore the issue of how to assign specific wavelengths to lightpaths. Based on the grooming and routing, we can then derive wavelength capacity requirements for all lightpaths. Similar to the GRWA problem, we formulate the GR problem as an ILP problem. The size of the GR ILP problem is much smaller than its corresponding GRWA ILP problem. Furthermore, we can

significantly improve the computational efficiency for the GR ILP problem by relaxing some of its integer constraints, which usually leads to approximate solutions for the GR problem. Once we solve the GR problem, we can then consider the WA problem, in which our goal is to derive a feasible wavelength assignment solution that assigns specific wavelengths to lightpaths based on their capacity requirements derived in the GR problem.

It is obvious that, in general, the decomposition method would not yield the optimal solution for the GRWA problem. However, we will provide a sufficient condition under which we show that the decomposition method does produce an optimal solution for the GRWA problem. We also develop a simple algorithm that finds a wavelength assignment solution under this sufficient condition.

9.3.1 The GR Problem

Let $t = [t_l]_{l \in L}$, a column vector containing lightpath capacity decision variables, where $t_l = \sum_{w \in W} y_{l,w}$ is the number of wavelengths needed for lightpath $l \in L$. Then, the GR problem can be formulated as:

$$\begin{aligned} \min \quad & \sum_{l \in L} t_l \\ \text{s.t.} \quad & Ax_d = u_d \quad d \in D \end{aligned} \quad (9.4)$$

$$Bt \leq |W|\mathbf{1} \quad (9.5)$$

$$\sum_{d \in D} s_d x_{l,d} \leq gt_l \quad l \in L \quad (9.6)$$

x binary variable and t integer variable.

We refer the above ILP problem as the GR ILP problem. We now present the following result:

Proposition 1 *If x_d and y_w are feasible solutions for the GRWA ILP problem, then x_d and t are feasible solutions for the GR ILP problem, where $t = \sum_{w \in W} y_w$.*

Proof. We first note that by summing over $w \in W$ in (9.2) it leads to (9.5). Secondly, (9.3) is the same as (9.6). Hence, the result follows. \square

Based on Proposition 1, we have

Proposition 2 *If x_d^* and t^* are the optimal solutions of the GR ILP problem, and there exists a binary y_w^* such that $\sum_{w \in W} y_w^* = t^*$ and $By_w^* \leq \mathbf{1}$ for $w \in W$, then x_d^* and y_w^* are the optimal solutions of the GRWA ILP problem.*

Proof. Suppose x_d and y_w are feasible solutions for the GRWA ILP problem, then based on Proposition 1, x_d and $t = \sum_{w \in W} y_w$ are feasible solutions for the GR ILP problem. Since x_d^* and t^* are the optimal solutions of the GR ILP problem, we have $\sum_{w \in W, l \in L} y_{l,w}^* = \sum_{l \in L} t_l^* \leq \sum_{l \in L} t_l = \sum_{w \in W, l \in L} y_{l,w}$. Therefore, the conclusion follows. \square

Obviously, the GR ILP problem is much easier to solve than the GRWA ILP problem since it has fewer integer variables and fewer constraints (e.g., see numerical examples in Section 9.4). More importantly, we can now relax the integer constraint on t in the GR ILP problem and solve a relaxed mixed ILP problem and then round up the values of t to obtain a solution for the GR problem. This would dramatically improve the computational efficiency. On the other hand, the relaxation approach is much less effective for the GRWA ILP problem since all its decision variables are binary. In general, if most lightpaths have relatively high wavelength counts (i.e., the values of their corresponding components in t are large), then the relaxed GR ILP problem often produces very good solutions for the GR problem, as illustrated by our numerical examples in Section 9.4. This is simply because if the optimal value of t_l is large, then the error of rounding up is relatively small.

9.3.2 The WA Problem

The WA problem of our interest is to find a binary solution y such that

$$\sum_{w \in W} y_w = t \quad \text{and} \quad By_w \leq \mathbf{1} \quad \text{for } w \in W,$$

where t is a feasible (or optimal) solution of the GR problem. This problem can be viewed as an ILP problem (without an objective function), which is much easier to solve than the GRWA ILP and the (relaxed) GR ILP problems. For example, it can be solved for networks with a few hundred nodes and lightpaths in seconds or minutes by using commercially available LP software, e.g., CPLEX. Based on Proposition 2, we know that if x and t are optimal solutions of the GR problem and the WA problem has a feasible solution y , then x and y are optimal solutions of the GRWA problem. In case when we cannot find a feasible solution for the WA problem, we can either increase the number of wavelengths in W in the WA problem (note that we can always find a feasible solution for the WA problem if W has enough wavelengths), or we can use $W^* \subset W$ in the GR problem (specifically, replace $|W|$ with $|W^*|$ in (9.5)) but still use W in the WA problem. Obviously, the latter approach is preferred in which case the decomposition method provides a feasible solution for the GRWA problem. An alternative approach is to use wavelength conversion via lightpath regeneration, which is equivalent to modifying L by breaking some lightpaths into two or more lightpaths. In addition, there are other possible remedies available to alleviate the infeasibility of the WA problem.

Though the WA problem can be solved as an ILP problem, it is also possible to solve it directly based on some heuristic algorithms (e.g., see [3]). In what follows, we consider a special type of the GRWA problem, in which the lightpaths satisfy a certain condition. Under such a condition, we show that a feasible solution for the corresponding WA problem can always be found, and we also develop an algorithm for finding a feasible solution. Without loss of generality, we assume that the capacity of every lightpath is one wavelength (i.e., $t_l = 1$ for every $l \in L$). For a lightpath whose capacity is more than one wavelength, we can treat it as several identical parallel lightpaths, each of which has capacity of one wavelength. Let p_e ($e \in E$) be the number of lightpaths that traverse fiber span e , and $p = \max_{e \in E} p_e$, which is the minimum number of wavelengths required for the network.

Define:

$$E_l := \{e \in E \mid e \text{ is on lightpath } l\}, l \in L;$$

$$L_e := \{l \in L \mid l \text{ traverses fiber span } e\}, e \in E.$$

We now present the following algorithm for the WA problem.

Algorithm 1 (for the WA problem)

1. Select an initial lightpath $l_0 \in L$ (arbitrarily), and assign a wavelength to l_0 .
2. Suppose $E_{l_0} = \{e_1, e_2, \dots, e_k\}$. Set $L_0 = \{l_0\}$. For $i = 1$ to k , do
 - a) Assign a wavelength to every lightpath $l \in L_{e_i} \setminus \cup_{0 \leq j < i} L_j$ such that no two lightpaths in L_{e_i} share the same wavelength (note that $L_{e_i} \setminus \cup_{0 \leq j < i} L_j$ is a subset of lightpaths in L_{e_i} to which wavelengths have not been assigned yet).
 - b) Let

$$L_i = L_{e_i} \setminus \cup_{0 \leq j < i} L_j,$$

$$E_i = \cup_{l \in L_i} E_l \setminus \{e_i\}.$$

We note that L_i is the set of lightpaths to which wavelengths are assigned in Step 2(a) and E_i is the set of fiber spans that are on at least one lightpath in L_i (excluding fiber span e_i).

3. For $i = 1, 2, \dots, k$, apply the procedure in Step 2 to E_i (with E_{l_0} being replaced with E_i), and continue until all the lightpaths in L are assigned (note that, since all the fiber spans in E_{l_0} have been considered already in Step 2, we can simply replace E_i by $E_i \setminus E_{l_0}$).

To study some useful properties associated with Algorithm 1, we first introduce the following terminologies:

Definition

1. We say a lightpath l and a fiber span e are connected (via fiber spans $\{e_1, \dots, e_m\}$ and lightpaths $\{l_1, \dots, l_m\}$) if there exist a set of fiber spans $\{e_1, \dots, e_m\}$ and a set of lightpaths $\{l_1, \dots, l_m\}$ such that $e_i \in E_{l_{i-1}}$ for $i = 1, \dots, m+1$ (where $l_0 \equiv l$ and $e_{m+1} \equiv e$) and $l_i \in L_{e_i}$ for $i = 1, \dots, m$.

2. We say two lightpaths l_0 and l_m are connected (via fiber spans $\{e_1, \dots, e_m\}$ and lightpaths $\{l_1, \dots, l_{m-1}\}$) if there exist a set of fiber spans $\{e_1, \dots, e_m\}$ and a sequence of lightpaths $\{l_1, \dots, l_{m-1}\}$ such that $e_i \in E_{l_{i-1}}$ and $l_i \in L_{e_i}$ for $i = 1, \dots, m$.
3. We say a set of lightpaths $\{l_1, \dots, l_m\}$ is a lightpath cycle if $E_{l_i} \cap E_{l_{i+1}} \neq \emptyset$ (i.e., lightpaths l_i and l_{i+1} share at least one common fiber span) for $i = 1, \dots, m$ ($l_{m+1} \equiv l_1$).
4. We say a lightpath cycle $\{l_1, \dots, l_m\}$ is a complete lightpath cycle if $E_{l_1} = \dots = E_{l_m}$, otherwise it is a non-complete lightpath cycle.

To help understand what is a lightpath cycle, consider the network depicted in Fig. 9.1. The network has four nodes (A, B, C, D), three fiber spans (A–B, B–C, B–D), and three lightpaths (A–B–C, C–B–D, D–B–A). It is clear that the three lightpaths (A–B–C, C–B–D, D–B–A) constitute a lightpath cycle, however it is a non-complete cycle.

We now present the following properties associated with Algorithm 1.

Proposition 3

1. Every fiber span in E_i is on at least one lightpath in L_i ;
2. For $1 \leq j \leq i$, $e_j \notin E_i$;
3. $L_i \cap L_j = \emptyset$ ($i \neq j$);
4. If $l \in L_i$, then it does not traverse fiber spans $\{e_1, \dots, e_{i-1}\}$;
5. If $E_i \cap E_j \neq \emptyset$ ($i \neq j$), then there exists a lightpath cycle with one lightpath in L_i and one lightpath in L_j ;
6. If a lightpath in L_i is connected to another lightpath in L_j in two different ways via lighpaths in $L \setminus \cup_{0 \leq h \leq k} L_h$ and fiber spans in $E \setminus E_{l_0}$, then there exists a lightpath cycle $\{l_1, \dots, l_m\}$ such that $E_i \cap (E_{l_{i_1}} \cap E_{l_{i_1+1}}) \neq \emptyset$ and $E_j \cap (E_{l_{i_2}} \cap E_{l_{i_2+1}}) \neq \emptyset$, where $1 \leq i_1 < i_2 \leq m$.

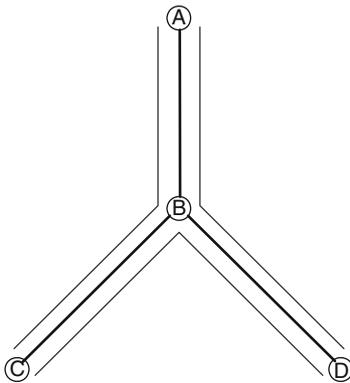


Fig. 9.1. A 4-Node network

Proof. We want to reiterate the fact that wavelengths are assigned to lightpaths in $L_i \subset L_{e_i}$ in Step 2(a).

1. By definition.
2. By definition, $e_i \notin E_i$. For $1 \leq j < i$ and $e \in E_i$, it is clear that we have assigned wavelengths to all the lightpaths in L_{e_j} before Step 2(a) while at least one lightpath in L_e has not been assigned by a wavelength before Step 2(a). Hence, $e_j \notin E_i$.
3. All the lightpaths in L_j are assigned by wavelengths at the end of Step 2(a) and they will not be considered again in later iterations.
4. By the same argument as in (2).
5. Suppose $e \in E_i \cap E_j$. Based on (1), e is on one lightpath in L_i , say l_i , and on another lightpath in L_j , say l_j . Furthermore, l_i and l_j traverse e_i and e_j , respectively, which are both on lightpath l_0 . Therefore, we have a lightpath cycle $\{l_i, l_0, l_j\}$.
6. The same argument used in (5) can be applied here as well. \square

In general, one needs to be careful about what wavelengths to use in Step 2(a) of Algorithm 1, otherwise it is possible that it may not produce a feasible solution for the WA problem. For example, consider the following example in which $E_{l_0} = \{e_1, e_2\}$, $L_1 = \{l_1\}$, $L_2 = \{l_2\}$, and $E_1 = E_2 = \{e\}$. If we assign the same wavelength to l_1 and l_2 , then we end up with assigning one wavelength to l_1 and l_2 on fiber span e , which is not permissible. Therefore, we have to assign u_1 and u_2 with different wavelengths.

It is clear that the number of different wavelengths needed in the WA problem is at least p . In what follows, we provide a sufficient condition under which p different wavelengths are enough to solve the WA problem.

Theorem 1 *If a network does not contain any non-complete lightpath cycle, then Algorithm 1 can produce a feasible solution for the WA problem which only needs p wavelengths.*

Proof. Since the network does not contain any non-complete lightpath cycle, based on (4), (5), and (6) in Proposition 3, we have (i) $E_i \cap E_j = \emptyset$ and (ii) no lightpath in L_i is connected to lightpath in L_j ($i \neq j$). Hence, when doing wavelength assignment for lightpaths in L_i in Step 2(a) we can use arbitrary wavelengths, and it guarantees that it is permissible (i.e., no two lightpaths that traverse the same fiber span would be assigned to the same wavelength). By repeating this argument, we can show that, in Algorithm 1, we can use arbitrary wavelengths in Step 2(a) and obtain a feasible solution for the WA problem. Since wavelengths used in Step 2(a) can be arbitrary, the maximum number of different wavelengths needed throughout Algorithm 1 should be no more than p . This completes our proof. \square

Theorem 1 implies that if a network does not contain any non-complete lightpath cycle, we can find a solution for the WA problem which only needs p wavelengths. In [3], the problem of whether the WA problem can be solved

with p wavelengths was also studied. However, we believe that the result there (Theorem 2 in [3]) is incorrect, which states that if a network is acyclic then its WA problem can be solved with p wavelengths. The network in Fig. 9.1 is a counter-example to this result. It is a tree (hence acyclic). Clearly we have $p = 2$, but need three wavelengths for its WA problem.

Since t in the WA problem is a feasible solution for the GR problem, i.e., $Bt \leq |W|\mathbf{1}$, we have $p \leq |W|$. This, together with Theorem 1, leads to the following result:

Theorem 2 *If a network does not contain any non-complete lightpath cycle, Algorithm 1 produces a feasible solution for the WA problem, and the decomposition method gives an optimal solution for the GRWA problem.*

To test whether a network contains any non-complete lightpath cycle, one can obviously use the exhaustive search method: finding all lightpath cycles and then test if any of them is non-complete. Clearly the complexity of this exhaustive search method grows exponentially. Currently, we do not have an efficient method to verify if a network contains any non-complete lightpath cycle. In fact, this problem itself could be NP-complete, just like the wavelength assignment problem.

In the case that a network contains non-complete lightpath cycles, let c^* be the minimum number of lightpaths that need to be removed from the network so that the remaining portion of the network does not contain any non-complete lightpath cycles. Then we have

Theorem 3 *There exists a feasible solution for the WA problem which requires at most $c^* + p$ wavelengths. Therefore, if $c^* + p \leq |W|$, then we can find a feasible solution for the WA problem and the decomposition method still gives an optimal solution for the GRWA problem.*

Before closing this section, we should point out that if the result in Theorem 3 can be further refined, then it can lead to better upper bounds on the number of wavelengths required for the WA problem.

9.4 Numerical Results

In this section, we present four sets of numerical examples. All ILPs and mixed ILPs were solved by using CPLEX 7.0 on a Dell Precision 420 PC with two 1GHz processors. We compare the numerical results obtained based on the three methods proposed in the previous two sections: the GRWA ILP formulation, the decomposition method combined with the GR ILP formulation, and the decomposition method combined with the relaxed GR ILP formulation. The run time for the decomposition method includes the run times for both the (relaxed) GR ILP problem and the WA problem. The run time for the WA problem in all four examples is very fast (it is less than a second in the

first three cases and less than 3 seconds in the last case). Our numerical results clearly indicate that the decomposition method combined with the relaxed GR ILP formulation produces quite good results with reasonably small run times.

Example 1. This is a relatively small network with 12 nodes, 17 fiber spans, 24 lightpaths, and 104 traffic demands (with different sizes). For this example, we were able to obtain the optimal solution based on the GRWA ILP formulation. The results are presented in Table 9.1.

Table 9.1. Numerical results for Example 1

	Run Time	Solution
GRWA ILP	400 seconds	128
GR ILP	80 seconds	128
Relaxed GR ILP	2 seconds	136

Example 2. The network we consider in this example has 30 nodes, 38 fiber spans, 47 lightpaths, and 242 demands (with different sizes). The results are presented in Table 9.2. For the GRWA ILP problem, we stopped the CPLEX program after 75 hours and obtained a feasible solution with objective value 249.

Table 9.2. Numerical results for Example 2

	Run Time	Solution
GRWA ILP	>75 hours	249
GR ILP	37 hours	189
Relaxed GR ILP	12 seconds	202

Example 3. The network in this example has 49 nodes, 75 fiber spans, 155 lightpaths, and 238 demands (with different sizes). It is a medium size network. For this example, the decomposition method based on the relaxed GR ILP problem produced a solution with value 345 in about 13 minutes, while the CPLEX program did not even return a feasible solution for the GRWA ILP and GR ILP problems after 40 hours (at which point we stopped the program). From the CPLEX program, we were also able to obtain a lower bound (based on the GR ILP problem) 328 for the objective function. Hence, the solution provided by the relaxed GR ILP-based decomposition method is within 5% of the lower bound. We note that the WA problem was solved in 0.37 seconds for this example. The results are presented in Table 9.3.

Table 9.3. Numerical results for Example 3

	Run Time	Solution
GRWA ILP	>40 hours	No Solution
GR ILP	>40 hours	No Solution
Relaxed GR ILP	13 minutes	345

Example 4. The network in this example has 144 nodes, 162 fiber spans, 299 lightpaths, and 600 demands (with different sizes). It is a relatively large network (a typical size for a nation-wide network). For this example, the method based on the relaxed GR ILP problem produced a solution in about 38 minutes, and the CPLEX program did not even return a feasible solution for the GRWA ILP and GR ILP problems after 100 hours (at which point we stopped the program). The WA problem in this case was solved in 2.67 seconds for this example. The results are shown in Table 9.4.

Table 9.4. Numerical results for Example 4

	Run Time	Solution
GRWA ILP	>100 hours	No Solution
GR ILP	>100 hours	No Solution
Relaxed GR ILP	38 minutes	431

References

1. A. Aggarwal, A. Bar-Noy, D. Coppersmith, R. Ramaswami, B. Schieber, and M. Sudan, "Efficient routing and scheduling algorithms for optical networks," In *Proceedings of the 5th ACM-SIAM Symposium on Discrete Algorithms*, pp. 412–423, Jan. 1994.
2. V. Auletta, I. Caragiannis, C. Kaklamanis, and P. Persiano, "Random path coloring on binary trees," In *Proceedings of the 3rd International Workshop on Approximation Algorithms for Combinatorial Optimization Problems (APPROX'00)*, LNCS 1913, Springer, pp. 60–71, 2000.
3. I. Chlamtac, A. Ganz, and G. Karmi, "Lightpath communications: an approach to high bandwidth optical WAN's," *IEEE Transactions on Communications*, Vol. 40, No. 7, pp. 1171–1182, 1992.
4. L. A. Cox, J. Sanchez, and L. Lu, "Cost savings from optimized packing and grooming of optical circuits: mesh versus ring comparisons," *Optical Networks Magazines*, pp. 72–90, May/June, 2001.
5. Z. Ding and M. Hamdi, "Clustering techniques for traffic grooming in optical WDM mesh networks," *Proceedings of 2002 Global Telecommunications Conference*, Vol. 3, pp. 2711–2715, 2002.

6. W. Grover, J. Doucette, M. Clouqueur, D. Leung, and D. Stamatelakis, "New options and insights for survivable transport networks," *IEEE Communications Magazine*, Vol. 40, No. 1, pp. 34–41, January 2002.
7. J. Q. Hu and B. Leida, "Traffic Grooming, Routing, and Wavelength Assignment in Optical WDM Mesh Networks," *Proceedings of 2004 INFOCOM*, pp. 495–501, 2004.
8. C. Lee and E. K. Park, "A genetic algorithm for traffic grooming in all-optical mesh networks," *Proceedings of 2002 IEEE International Conference on Systems, Man and Cybernetics*, Vol. 7, 2003.
9. M. Mihail, C. Kaklamanis, and S. Rao, "Efficient access to optical bandwidth–wavelength routing on directed fiber trees, rings, and trees of rings," In *Proceedings of the 36th Annual Symposium on Foundations of Computer Science*, pp. 548–557, Milwaukee, Wisconsin, 23–25 October, 1995.
10. K. G. Patrocínio Jr. and G. R. Mateus, "A lagrangian-based heuristic for traffic grooming in WDM optical networks," *Proceedings of 2003 Global Telecommunications Conference*, Vol. 5, pp. 2767–2771, 2003.
11. P. Raghavan and E. Upfal, "Efficient routing in all-optical networks," In *Proceedings of the 26th ACM Symposium on Theory of Computing*, pp. 134–143, May 1994.
12. R. Ramaswami and K. Sivarajan, *Optical Networks: A Practical Perspective*, Morgan Kaufmann Publishers, 1998.
13. S. Subramaniam, M. Azizoglu, and A. Somani, "On optimal converter placement in wavelength-routed networks," *IEEE/ACM Transactions on Networking*, Vol. 7, No. 5, pp. 754–766, 1999.
14. Y. Tao, *Design and Performance Analysis of WDM Optical Networks*, Ph.D. Dissertation, Boston University, 2006.
15. Y. Tao and J.Q. Hu, "Amplifier Placement for Optical Networks," Manuscript, Department of Manufacturing Engineering, Boston University, 2003.
16. K. Zhu and B. Mukherjee, "Traffic grooming in an optical WDM mesh network," *Proceedings of IEEE International Conference on Communications*, pp. 721–725, Helsinki, Finland, June, 2001.
17. "The intelligent optical core: improving bandwidth efficiency," Sycamore Networks, white paper, 2001.

Survivable Traffic Grooming

Arun K. Somani

10.1 Introduction

Optical fiber medium using wavelength division multiplexing (WDM) offers tremendous transmission bandwidth to deliver high-bandwidth services cost effectively. A WDM-based network divides the fiber capacity into non-overlapping wavelength channels, each of which operates at a transmission rate compatible with electronics. The routing function is controlled by the optical layer management planes. The nodes equipped with *optical cross-connects* support routing by switching different wavelengths at input ports to different output ports. The switching is managed by the control plane. An end-to-end path, called a *lightpath*, is established using the same wavelength on all links on the chosen path, or a different wavelength on a different link on the path with required *wavelength conversion* provided by intermediate links on the path. An all optical domain on path can provide complete transparency in data transmission. A network layer, called the WDM layer, in a layered architecture provides interfaces to other layers such as SONET/SDH gear, IP (Internet Protocol) ATM (Asynchronous Transfer Mode) or any other transport technology to connect to a lightpath.

10.1.1 Protection and Restoration

As wavelength routing paves the way for network throughput of possibly hundreds of Tb/s, network survivability assumes critical importance. A loss or damage to a fiber is a common means of a greater loss. A short network outage can lead to huge data loss. Thus a connection being carried in the network also needs high protection. Survivability refers to the ability of the network to reconfigure and reestablish communication upon failures. According to industry standards, the expected availability requirements are 99.99% or higher. The basic types of network failures generally considered are link and node failure. Cable cuts that cause link failures are common in optical networks. A node failure is usually due to equipment failure at the node. Channel failure

is unique to WDM networks. A channel failure is usually caused by the failure of transmitting or receiving equipment operating on that channel.

Several survivability techniques have been proposed [1, 2, 3] and can be classified into two general categories: preplanned protection and dynamic restoration. In preplanned protection-based techniques resources are already planned, typically at the time of establishing a connection, to recover from network failures and hence recovery is faster. During the operation phase these reserved resources remain idle. Upon the occurrence of failure, reserved resources are used to recover from the failure according to protection protocols. In contrast, in dynamic restoration, the resources used for recovery from failure are not reserved at the time of connection establishment, but are discovered dynamically when a failure occurs. As is obvious, dynamic restoration uses resources efficiently, but the restoration time is usually longer. Moreover, 100% service recovery cannot be guaranteed as it is not guaranteed that the spare capacity is available at the time of failure. To guarantee the service, a preplanned protection is the preferred approach. Most researchers therefore focus on preplanned protection. Figure 10.1 shows a classification of survivable design techniques that can be deployed at various layers.

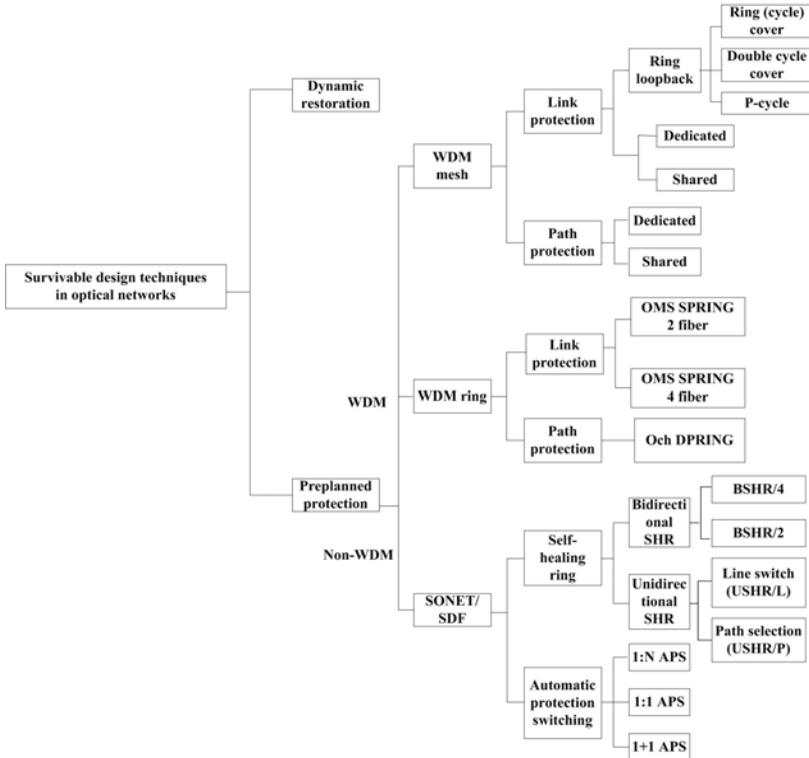


Fig. 10.1. Survivable design techniques in optical networks

Since the protection can be provided at many layers, a relevant question is, “which layer is the best?”. Although there is no consensus, most researchers believe that the optical layer protection is the best. It is very efficient in handling failures such as fiber cuts, which are probably the biggest cause of failure. To provide protection at any other layer requires that the layer must keep track of all logical links that may be affected by a single fiber failure. Notice that a logical path may use many lightpaths and they may be routed on a fiber/link. This implies that the relationship between established routes and fiber links must be collected and managed. This imposes additional overheads on protocols, algorithms, and processing during the normal operation. It is also possible that such an information cannot be made available by the other service provider. The survivability of managing layer, upon detection of a failure, may need to establish additional logical links upon failure that may lead to delay. Moreover, preplanning may require establishing logical links and/or lightpaths that may or may not use certain fiber links. Thus, in planning for a failure, the management, and actual recovery may overwhelm the management system. Moreover, optical layer protection and restoration provides an additional resilience in the network. For example, optical layer can be used to provide resilience against multiple failures. Furthermore, the layers above the optical layer may not be fully able to provide the protection. If protection is provided at many layers, then significant cost savings can be obtained by making use of optical layer protection and restoration.

It should, however, be clear that optical layer protection cannot handle any faults in the higher layer of network. Because of the property of protocol transparency, the optical layer may be unaware of what exactly is carried on the lightpaths, and therefore, cannot monitor the traffic to sense any degradation. A protection mechanism must also consider multiple, in particular two, link failure scenarios due to common routing of fiber cables.

10.2 WDM Mesh Network Protection

Survivable design in optical mesh networks involve high efficiency in capacity utilization and fast restoration. The problem is complex because there are multiple routes that can be used to recover from failures. A path that is used to carry the traffic under normal conditions is called a *primary* or a *working* path. A path that is reserved for protection is called a *backup* or a *protection* path. The problem becomes more complex when each wavelength is also groomed with smaller traffic streams as the number of affected traffic streams due to a single link failure is much larger than a wavelength-based connection. Protection schemes in mesh networks are broadly classified as either link-based or path-based, as shown in Fig. 10.2.

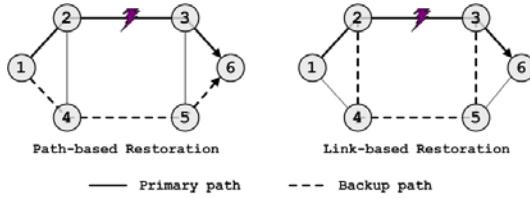


Fig. 10.2. Link- and path-based protection

A path-based method employs end-to-end detouring. The traffic is rerouted by the source node of affected path to the destination. A pair of link-disjoint paths are established between the source and destination nodes of each connection request. One path is used as the working path, and the other is used as the backup path. When a link failure occurs, the source node of the affected path switches the traffic to its backup path. In link-based protection, backup paths for each link are pre-computed. Upon the failure of a link, the working connections on the link are switched by the two end nodes of the link to their corresponding backup paths. Thus, a link-based method employs local detouring and the traffic is rerouted around the failed link.

An alternate link based protection scheme is called a subgraph-based routing [4]. The idea behind the *subgraph-based routing* scheme is to plan network resource utilization in such a way that, for any link failure, there exists an alternate path for every accepted request. When a link fails, all paths that get affected by the link failure are reassigned to their new paths.

The link- and path-based protection schemes can use either dedicated or shared resources. Dedicated resources are exclusively reserved for backup paths. In case of dedicated protection, the optical cross-connects (OXC) for the backup path can be preconfigured requiring only the source node to switch the path. If two or more primary paths are link disjoint, then the corresponding backup paths can share resources as only one of the primary paths will fail due to a single link failure. This is called backup multiplexing. This utilizes fewer resources, but requires a more complicated management. In shared backup path-based protection (SBPP), it is necessary to configure OXC for the backup path accordingly after a link failure occurs. This inevitably increases the recovery delay. SBPP tends to use lesser total capacity than shared link-based protection.

The link-based methods limits the choices for alternatives. However, link-based protection tends to be faster than shared path-based protection because only the two end nodes of the failed link are involved in recovery. One drawback for link-path protection in WDM networks, however, is that the backup path must necessarily use the same wavelength as the primary path. To tolerate node failures, a link based scheme can be extended to a two-link segment-based protection scheme so that the path can be routed around a failed node.

10.2.1 Alternate Link-Based Protection Schemes

A few special types of link-based protection schemes have also been developed, which aim to benefit by using the fast restoration of ring-like protection or save resources. Notable among these are: *Double Cycle Cover*, the *p-Cycle*, and subgraph-based routing schemes.

In the double cycle cover [5], the network is represented by a directed graph. A set of cycles are embedded on the given topology. Each link is covered by two directed cycles, one in each direction. For planar graphs, the required set of protection cycles can be found in polynomial time, but no known polynomial-time algorithm for non-planar graphs is known [5]. The model uses fiber-based recovery, similar to the SONET BLSR ring. On each link, half of the capacity is reserved for backup and the other half is used for working traffic in each direction to protect the traffic in opposite direction. Since the protection switches can be preconfigured, this method can achieve fast restoration, but at the cost of 100% redundancy.

The *p-Cycle* protection method [6, 7] also uses a cyclic layout of spare capacity to provide protection. When a link fails, only the nodes neighboring the failure need to perform real-time switching. The key difference between the *p-Cycle* protection and the ring cycle protection, such as double cycle cover, is that the *p-Cycle* protection not only protects the links on the cycle, as in the ring protection, it also protects straddling links. A straddling link is an off-cycle link whose two end nodes are both on the cycle. This important property effectively improves the capacity efficiency of *p-Cycles*.

Figure 10.3 depicts an example that illustrates *p-Cycle* protection. In Fig. 10.3 (a), A-B-C-D-E-A is a *p-Cycle* formed using spare capacity. When an on-cycle link A-B fails, the *p-Cycle* provides protection as shown in Figure 10.3(b) for traffic from A to B by using the path A-E-D-C-B. A reverse path is followed for traffic in the other direction, i.e., path B-C-D-E-A is used. When a straddling link B-D fails, the *p-Cycle* provides two backup paths, namely B-A-E-D and B-C-D for traffic from B to D as shown in Fig. 10.3(c). Thus the full capacity of a straddling link can carry working traffic, as each backup path is able to provide half of the capacity. For traffic in the other direction, i.e., D to B, paths are used in the reverse direction. Thus straddling links are covered fully by a cycle whereas an on-cycle link can use only half the capacity for primary connections, the other half has to be reserved for protection.

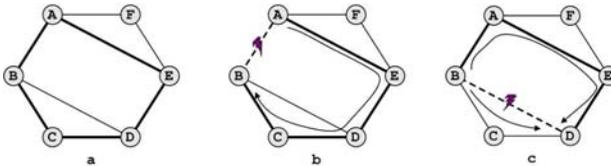


Fig. 10.3. (a) A *p-Cycle*. (b) protecting on-cycle link. (c) protecting straddling link

The subgraph-based routing strategy [4] provides a passive form of redundancy in the event of a fault occurrence from a given set of faults. The routing algorithm maintains $|F| + 1$ copies of states of the network resource utilization where $|F|$ is the number of possible faults. The i th copy of the state, S_i , represents the state of network when all accepted requests are routed without using the failed resources due to the presence of fault f_i . One copy of the network state, S_0 , represents a fault free state. When a request arrives, it is accepted only if it can be routed in each of $|F| + 1$ states using the free resources in the corresponding state. If a request is accepted, the resources used in each state are marked as busy in that state. If a fault f_i occurs, the state S_i provides the resources that must be used to route all accepted requests.

The scheme is passive in the sense that resources are available, but not reserved or preconfigured at the time of setting up connections. An end user experiences nominal interruption in service when a fault occurs and the connections are rerouted/restored. A system process to carry out restoration is described in [8]. The key characteristics of this scheme are that no additional resources are used to provide connection redundancy, although a 100% guarantee against all known faults is provided using a path-based fault tolerance strategy for anticipated failures. It is a path based recovery strategy because it does not guarantee that any of the same links are used to reroute a connection upon the occurrence of a fault. A disadvantage is that the scheme can potentially require a complete reconfiguration of the network to a predetermined new state. However, it is shown to be most resource efficient [4].

10.2.2 Dynamic Traffic Handling

Dynamically provisioning connections on demand is becoming more important in groomed transport networks. Early research has mostly focused on static traffic. Dynamic traffic can be handled by either using a strategy where the primary and backup paths are determined when a request arrives or backup resources are reserved and a connection is accepted only if both primary and backup paths can be found among the respective available resources for each. The objective of partitioning the resources is to guarantee that the capacity available for routing randomly arriving connection requests will be 100% protected by the reserved protection capacity. A p -Cycle based scheme can also deal with dynamic traffic using the latter scheme of resource reservation and a two-step approach. The basic idea is to provision the network resources in two parts: protection resources and working-capacity resources in the p -Cycles. The design must also ensure that the p -Cycles are preconfigured. In the first step, one can find a set of p -Cycles to cover the network and reserve enough capacity in p -Cycles. In the second step, one routes the requests as they arrive. The reserved capacity on p -Cycles leads to less control signaling overhead and less dynamic state information needs to be maintained to achieve protection. Thus, the p -Cycle based design has the advantage of fast recovery, less control signaling, and less dynamic state information to be maintained.

10.3 Protection Design: Optimization and Performance

The problem of restorable network design for a static traffic demand has been dealt with in [1, 2, 3, 9, 10, 11] and for dynamic traffic scenario in [12, 13, 14, 15, 16]. For protection mechanism design and performance analysis, a network is represented using a directed graph with N nodes and L links. For dedicated and shared path protection design, a set of K link-disjoint alternate routes for every source–destination pairs can be pre-computed. It has been shown that $K = 2$ or 3 is sufficient to achieve good performance to tolerate a single link failure. For a network with W wavelengths and no wavelength conversion, K link-disjoint alternate routes can be viewed as $K \times W$ paths, each of which is wavelength continuous path. In shared path protection scheme, the capacity optimization problem is to choose a pair of link-disjoint paths to serve as primary and backup for each connection request. The objective function is to minimize the total capacity required. The shared backup path protection (denoted as SBPP) can be implemented in two forms. In the first form, the goal is to minimize the total number of wavelengths used on all links when paths are not pre-configured. This is generalized to obtain the second scheme where the paths are pre-configured (denoted as PRE-SBPP).

An alternate approach is to use p -Cycle based protection (denoted as PCP), where the goal is to identify a set of simple distinct cycles that are sufficient to protect against all failures of interest. Longer cycles waste more capacity for redundancy and create long backup paths. A two-step algorithmic approach for the p -Cycle optimization, therefore, identifies a set of candidate p -Cycles in the first step and iteratively chooses a subset to deploy in actual implementations. Usually, the preselected set in the first step only includes cycles of up to a certain length. In a joint optimization design, one attempts to optimize the choice of routing working connections in conjunction with the p -Cycle selection, so that the total capacity is minimized.

In most cases, an Integer Linear Programming (ILP) formulation is used to choose optimal paths and/or cycles. We give an outline of ILP formulations for the three schemes, SBPP, PRE-SBPP, and PCP.

10.3.1 ILP Formulation

To formulate the ILP for the three cases, we use the following notations.

- $n = 1, 2, \dots, N$: Number assigned to each node
- $l, k = 1, 2, \dots, L$: Number assigned to each link
- $\lambda = 1, 2, \dots, W$: Number assigned to each wavelength
- $i, j = 1, 2, \dots, N(N - 1)$: Number assigned to s - d pair
- $K = 2$: Number of alternate routes between each s - d pair
- $p, r = 1, 2, \dots, KW$: Number assigned to a path
- (i, p) : Refers to the p th path for s - d pair i
- d_i : Demand for node pair i , in terms of number of lightpath requests
- ρ_l^n : It equals one if node n is an end node of link l , else zero (data)

The following notations are used for path information

Variable: $\delta^{i,p} = 1$ if (i, p) is chosen as a primary path, else 0

Variable: $\nu^{j,r} = 1$ if (j, r) is chosen as a restoration path, else 0

Data: $\epsilon_l^{i,p} = 1$ if link l is used in path (i, p) , else 0

Data: $\psi_\lambda^{i,p} = 1$ if λ is used by the path (i, p) , else 0

Variable: $g_{l,\lambda} = 1$ if λ is used by a restoration route that traverses link l

Variable: s_l : # of wavelengths used by backup lightpaths on link l

Variable: w_l : # of wavelengths used by primary lightpaths on link l

Data: $I_{(i,p)(j,r)} = 1$ if paths (i, p) and (j, r) share link. If $i = j$, then $p \neq r$

Data: $\Pi_{(i,p)(j,r)}^\lambda$: # of shared links between paths (i, p) and (j, r)

The following additional notations are used in formulation for PCP.

$c = 1, 2, \dots, P$: Number assigned to a cycle.

Data: $\omega_c^l = 1$ if link l is on cycle c ; else 0

Data: $\sigma_c^l = 1$ if link l is a straddling link on cycle c , else 0

Variable: $\tau_c^\lambda = 1$ if cycle c is chosen and uses wavelength λ , else 0

ILP Formulation for General Shared Path Protection

$$\text{OBJECTIVE :} \quad \text{Min} \sum_{l=1}^L (w_l + s_l) \quad (10.1)$$

1. *Link capacity constraint:*

$$w_l + s_l \leq W \quad 1 \leq l \leq L \quad (10.2)$$

2. *Demand constraint for each node pair:*

$$\sum_{p=1}^{KW} \delta^{i,p} = d_i \quad 1 \leq i \leq N(N-1) \quad (10.3)$$

3. *Primary link capacity constraint:* # of primary lightpaths on a link.

$$w_l = \sum_{i=1}^{N(N-1)} \sum_{p=1}^{KW} \delta^{i,p} \epsilon_l^{i,p} \quad 1 \leq l \leq L \quad (10.4)$$

4. *Spare capacity constraint:* Spare capacity required on link l .

$$s_l = \sum_{\lambda=1}^W g_{l,\lambda} \quad 1 \leq l \leq L \quad (10.5)$$

5. *Primary path wavelength usage constraint:*

$$\sum_{i=1}^{N(N-1)} \sum_{p=1}^{KW} \delta^{i,p} \epsilon_l^{i,p} \psi_\lambda^{i,p} + g_{l,\lambda} \leq 1 \quad 1 \leq l \leq L, 1 \leq \lambda \leq W \quad (10.6)$$

6. *Restoration path wavelength usage constraint:*

$$g_{l,\lambda} \leq \sum_{i=1}^{N(N-1)} \sum_{r=1}^{KW} \nu^{i,r} \epsilon_l^{i,r} \psi_\lambda^{i,r} \quad 1 \leq l \leq L, 1 \leq \lambda \leq W \quad (10.7)$$

$$N(N-1)KW g_{l,\lambda} \geq \sum_{i=1}^{N(N-1)} \sum_{r=1}^{KW} \nu^{i,r} \epsilon_l^{i,r} \psi_\lambda^{i,r} \quad 1 \leq l \leq L, 1 \leq \lambda \leq W \quad (10.8)$$

7. *Demand constraint for node pair i :* There is one restoration route for each primary call. Let $u \in \{0, 1\}$ and $v = 1 - u$:

$$\sum_{p=uW+1}^{(u+1)W} \delta^{i,p} = \sum_{r=vW+1}^{(v+1)W} \nu^{i,r} \quad 1 \leq i \leq N(N-1) \quad (10.9)$$

8. *Constraint for topology diversity of primary and backup paths:* Primary and backup paths should be link disjoint. Let $m, n \in \{0, 1\}$; $s = 1 - m$; and $t = 1 - n$. The primary path of a node pair can be any one of the two alternate paths for this node pair. We use m th path of node pair i as primary path for node pair i , n th path of node pair j as primary path of node pair j , s th path of node pair i as backup path of node pair i , and t th path of node pair j as backup path for node pair j . If $I_{(i,m)(j,n)} = 1$,

$$(\nu^{i,sW+\lambda} \psi_\lambda^{i,sW+\lambda} \epsilon_l^{i,s} + \nu^{j,tW+\lambda} \psi_\lambda^{j,tW+\lambda} \epsilon_l^{j,t}) I_{(i,m)(j,n)} \leq 1 \quad (10.10)$$

$$1 \leq i, j \leq N(N-1), 1 \leq \lambda \leq W, 1 \leq l \leq L$$

Pre-Cross-Connected Shared Path Protection

All other constraints for this formulation remain the same except that the pre-cross-connected protection requires an additional constraint. Suppose link l , k , and m share a common node n . If backup path b_1 uses link l and k , and backup path b_2 uses link l and m , then b_1 and b_2 cannot share the backup capacity on link l .

$$(\nu^{i,p} \epsilon_l^{i,p} \epsilon_k^{i,p} \psi_\lambda^{i,p} + \nu^{j,r} \epsilon_l^{j,r} \epsilon_m^{j,r} \psi_\lambda^{j,r}) \rho_l^n \rho_k^n \rho_m^n \leq 1 \quad (10.11)$$

$$1 \leq i, j \leq N(N-1), 1 \leq p, r \leq KW, 1 \leq l, k, m, \leq L, 1 \leq n \leq N$$

P-Cycle Protection

The objectives, link capacity constraint, demand constraint, and primary link capacity constraint remain as above. The other constraints are as follows.

Spare capacity constraint:

$$s_l = \sum_{c=1}^P \sum_{\lambda=1}^W \omega_c^l \tau_c^\lambda \quad 1 \leq l \leq L \tag{10.12}$$

Primary path wavelength usage constraint: Only one primary path can use a wavelength λ on link l ; no p -Cycle can use the same λ on link l .

$$\sum_{i=1}^{N(N-1)} \sum_{p=1}^{KW} \delta^{i,p} \epsilon_l^{i,p} \psi_\lambda^{i,p} + \sum_{c=1}^P \omega_c^l \tau_c^\lambda \leq 1 \tag{10.13}$$

$$1 \leq l \leq L, 1 \leq \lambda \leq W \tag{10.14}$$

Restoration guarantee constraint for link l : There are enough p -Cycles and wavelengths to recover the failure of link l .

$$\sum_{i=1}^{N(N-1)} \sum_{p=1}^{KW} \delta^{i,p} \epsilon_l^{i,p} \psi_\lambda^{i,p} \leq \sum_{c=1}^P \sigma_c^l \tau_c^\lambda \tag{10.15}$$

$$1 \leq l \leq L, 1 \leq \lambda \leq W \tag{10.16}$$

Performance of Different Schemes

The capacity performance of the three protection schemes has been studied by experimenting on six topologies [15]. Figure 10.4 on left shows a Pan-European COST239 network with 11 nodes and 26 links with an average node degree of 4.7 when all links are present. To study the effect of average nodal degree of a network on the performance of different schemes, three more topologies are created by manipulating the original COST239 network by deleting 5 (dotted), additional 4 (dashed), and additional 3 (dot-dashed) edges sequentially. These edges are progressively minimally used for a traffic matrix shown in Fig. 10.5 on the left. This results into four topologies. Two more topologies used in the experiments are a 11-node 22-link NJ-LATA network, which has average nodal degree 4, and a 14-node 21-link NSFNET, which has an average nodal degree 3, as shown in Fig. 10.4 in middle and right, respectively.

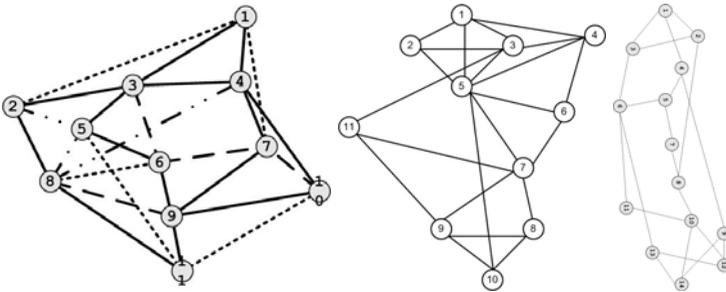


Fig. 10.4. 11N/26L COST 239, 11N/22L NJ LATA and 14N/21L NSFNET.

Node	1	2	3	4	5	6	7	8	9	10	11
1	0	1	1	3	1	1	1	1	1	1	1
2	1	0	5	8	4	1	1	10	3	2	3
3	1	5	0	8	4	1	1	5	3	1	2
4	3	8	8	0	6	2	2	11	11	9	9
5	1	4	4	6	0	1	1	6	6	1	2
6	1	1	1	2	1	0	1	1	1	1	1
7	1	1	1	2	1	1	0	1	1	1	1
8	1	10	5	11	6	1	1	0	6	2	5
9	1	3	3	11	6	1	1	6	0	3	6
10	1	2	1	9	1	1	1	2	3	0	3
11	1	3	2	9	2	1	1	5	6	3	0

Topology	Avr.		PRE-	
COST239	degee	SBPP	SBPP	PCP
26 links	4.7	816	948	794
Opt. Gap		$\leq 0.4\%$	$\leq 0.4\%$	$\leq 5.0\%$
21 links	3.8	899	1039	908
Opt. Gap		$\leq 3.0\%$	$\leq 2.0\%$	$\leq 3.0\%$
17 links	3.1	982	1108	1118
Opt. Gap		$\leq 3.0\%$	$\leq 2.0\%$	$\leq 0.4\%$
14 links	2.5	1280	1427	1717
Opt. Gap		$\leq 0.4\%$	$\leq 3.0\%$	$\leq 5.0\%$

Fig. 10.5. Traffic Matrix and Capacity used in terms of # of wavelength links

For the traffic matrix, the total capacity used by three different protection schemes in the four variations of COST 239 network are presented in Figure 10.5 on the right by solving the ILP formulation. As the average nodal degree increases gradually in the modified COST239 network from 14 links to 26 links, the total capacity required for establishing restorable connections for all the requests decreases for all three protection schemes. The decline of total capacity used is due to the decline of both working and backup capacity as the network connectivity increases. The improvement in the capacity utilization is because the alternate paths become shorter and the opportunity for backup capacity sharing in SBPP and PRE-SBPP increases as the network connectivity increases. The improvement of capacity efficiency with the increase in network connectivity is more dramatic in PCP scheme than in path-based protection (SBPP and PRE-SBPP), which is good for low-connectivity networks. The performance optimization gap in solving the ILP is also shown.

To further validate the reasoning and to study the effect of network connectivity on the capacity performance of three different protection schemes, two randomly generated traffic matrices with different characteristics are used to conduct experiments on the six network topologies. The number of requests in Type I traffic is smaller than that in Type II traffic. The average total capacity used for each protection scheme is the average value of ten sets of requests. Table 10.1 shows the average total used capacity by three different schemes in six topologies for Type I and Type II traffic, respectively.

The results indicate that general shared-path protection and pre-cross-connected shared-path protection use less total capacity than p -Cycle protection in low-connectivity networks, while they are comparable in high-connectivity networks. The p -Cycle protection scheme appears to be a better option for protection in high-connectivity networks. Pre-cross-connected shared-path protection uses more total capacity than general shared backup path protection. Pre-cross-connected shared-path protection can achieve faster restoration while remaining to be a path-based method.

Table 10.1. Average total capacity used by different protection schemes in six topologies for Type I and Type II traffic (number of wavelength links) for various topologies

Topology	Avr. degree	Type I Traffic			Type II Traffic		
		SBPP	PRE-SBPP	PCP	SBPP	PRE-SBPP	PCP
Cost239/26 links	4.7	72	74	70	101	106	93
NJ-LATA	4.0	78	81	83	111	118	114
Cost239/21 links	3.8	77	80	83	110	117	112
Cost239/17 links	3.1	91	97	110	130	142	155
NSFNET/21 link	3.0	105	113	132	147	162	181
Cost239/14 links	2.5	106	115	136	147	164	186

10.4 Traffic Grooming and Survivable Design

As noted in earlier chapters, grooming is the ability to share resources among multiple requests needing the network resource. This sharing is possible and even desirable because the resources under question are expensive, individual entities need only a fraction of the resource, and multiplexing allows the resource cost to be amortized over the number of users. Given the generic nature of the problem, grooming can be provided within a layer or across layers, such as optical grooming and electronic grooming. The optical layer is typically equipped with reconfigurable switching fabrics and the optical grooming technique used depends on the granularity and time scale of the switching functionality available in the optical layer. Optical switching can be done at various levels of granularity and time scales that include: waveband switching, wavelength switching, sub-wavelength time slot level switching, burst switching, flow switching on lightpaths, and packet switching.

Wavelength level switching techniques allow circuit switched sharing of a wavelength. When supplemented with some additional hardware an overlaid control protocol can allow statistical sharing of a wavelength as well. For example, in sub-wavelength time slot level switching, the wavelength is divided into T fixed time slots and the optical switching fabric is reconfigured for every k time slots ($k = 1..T$). Optical circuits are configured in the form of a path or a tree, and are categorized into four main classes:

1. Point to Point (P2P): this technique allows a lightwave circuit to aggregate traffic between the convenor node and the end node of the circuit.
2. Point to Multi-point (P2MP): this technique allows a lightwave circuit to aggregate traffic from a source to multiple destinations.
3. Multipoint to Point (MP2P): this technique allows a lightwave circuit to aggregate traffic from multiple sources to the same destination.
4. Multipoint to Multi-point (MP2MP): this technique allows a lightwave circuit to aggregate traffic from multiple sources to multiple destinations.

In all cases, the protection mechanisms are designed using the same general principles described earlier. The different traffic streams using the same primary path, however, may use very different backup paths in case of grooming. Also, a wavelength channel may carry both primary and backup capacity. These factors affect the overall utilization of resources and performance.

The study in [13] addressed the problem of dynamically establishing dependable low-rate traffic stream connections in WDM mesh networks with traffic grooming capabilities. To establish a dependable connection, they pre-compute link-disjoint primary and backup paths between the source and destination node and use backup multiplexing to reduce the overhead of backup traffic streams. Two schemes for grooming traffic streams onto wavelengths were proposed, namely *Mixed Primary-Backup Grooming Policy* (MGP) and *Segregated Primary-Backup Grooming Policy* (SGP). In SGP scheme, a wavelength is either used for primary paths or backup paths, but not both. MGP scheme does not put that constraint. The simulation results showed that SGP performs better in mesh networks and MGP performs better in ring networks.

In [17], the authors proposed three approaches, namely protection-at-lightpath (PAL) level, mixed protection-at-connection (MPAC) level, and separate protection-at-connection (SPAC) level, for grooming a connection request with shared protection. The authors concluded that when the lower bandwidth connections outnumber higher bandwidth connections, it is beneficial to groom working paths and backup paths separately, especially when the number of grooming ports is sufficient. Otherwise, protecting each specific lightpath achieves the best performance.

In the following, we investigate the problem of how to *groom* subwavelength level requests efficiently in mesh restorable WDM networks, and formulate the corresponding path selection and wavelength assignment problem as ILP optimization problems. Since the ILPs become too complex too soon, based on the same design principles, we also develop a heuristic algorithm for routing dynamic traffic in grooming networks. In some cases the network resources may not be sufficient to protect all traffic streams to the full extent to provide 100% protection. For such cases, our approaches of survivable grooming network design is also extended to partially protected WDM grooming networks.

10.4.1 Formulation of the Optimization Problem

We again use the same network model as in the previous section and with W wavelengths and $K = 2$ disjoint alternate paths for each s-d pairs, we have $W \times K$ networks, one each of a wavelength. The first W networks, numbered from 1 to W , contain the first alternate path for each s-d pair and the second W networks, numbered from $W + 1$ to $2W$, contain the second alternate path. Figure 10.6 (left) illustrates this layered model of a 6-node network with 3 wavelengths, 2 connections with each having 2 link disjoint alternate paths. Please note that if a path among network 1 to W is selected as a primary path for a request, its backup paths can only be selected from the network

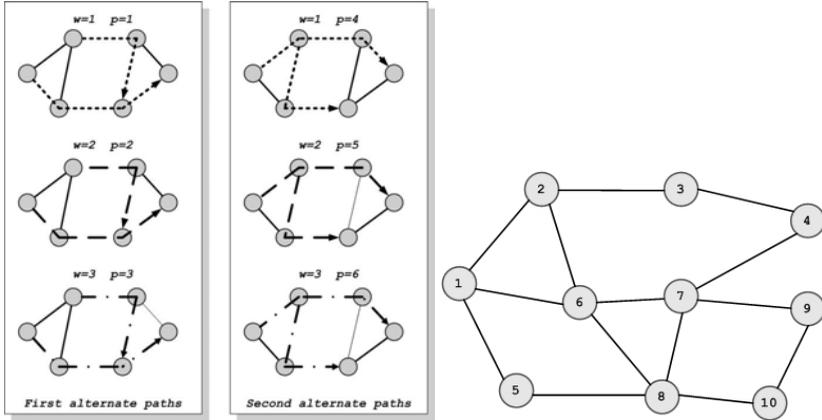


Fig. 10.6. (a) An example of layered network model with $W = 3$ and $K = 2$ and (b) an experimental topology

$W + 1$ to $2W$ in this layered network model. We consider 100% restoration guarantee for any single link failure for protected connections.

10.4.2 Shared Backup Reservation

In grooming WDM networks, the capacity reserved for restoration paths is more complicated. Let $B = \{b_1, b_2, \dots, b_k\}$ denotes the set of backup paths that traverse the wavelength w on link l . Let their respective capacities be $D = \{d_1, d_2, \dots, d_k\}$ and their respective primary paths be $P = \{p_1, p_2, \dots, p_k\}$. If none of the p_i 's have common links, the needed capacity on w is $\max(d_1, d_2, \dots, d_k)$. If some of the p_i 's have common links, their backup paths can still be groomed on wavelength w . However, the capacity to be reserved must be up to the summation of their capacities. The primary paths can be grouped according to their common links. Let $P^l = \{p_1^l, p_2^l, \dots, p_a^l\}$ denote the group of primary paths that have link l as their common link. The capacity required by this group for back up of link l is then given by $D^l = (d_1^l + d_2^l + \dots + d_a^l)$. It is possible that one primary path belongs to more than one group. The reserved capacity on wavelength w on link l is therefore the maximum value of the capacities required by all the groups, that is $D = \max(D^l)$.

10.4.3 Dedicated Backup Reservation

One simple and effective way of assigning backup capacities is to reserve dedicated capacity for each backup path. While choosing primary paths, instead of simply choosing the shortest path, we try to *minimize* the total *link-primary-sharing* (MLPS). The link-primary-sharing is defined as $s_l = \max(0, P_l - 1)$

where s_l denotes the link-primary-sharing of link l and P_l denotes the total number of primary paths that utilize link l . s_l can be viewed as the penalty assigned to link l when it is used by more than one primary path.

We formulate backup multiplexing and dedicated backup reservation schemes with MLPS goals as ILP optimization problems below.

To formulate the grooming survivable network design problem in a WDM mesh network with static traffic patterns as an ILP problem, we assume that the network is a single-fiber general mesh network. A connection request cannot be divided into several lower speed connection requests and routed separately from the source to the destination. The data traffic on a connection request should always follow the same route. The transceivers in a network node are fixed, hence the wavelength continuity constraint applies. In the following formulation, we assume that each grooming node has unlimited multiplexing and demultiplexing capability. This means that the network node can multiplex/demultiplex as many low-speed traffic streams to a lightpath as needed, as long as the aggregated traffic does not exceed the lightpath capacity. However, in practical systems, it may be a constraint.

The following additional parameters are needed for formulation over and above that defined in the last section.

C_l : Cost of using link l .

Variable $p_{l,w}^i$: =1 if wavelength w on link l carries primary, else 0.

Variable $r_{l,w}^i$: =1 if wavelength w on link l carries backup, else 0.

W_l : Number of wavelengths required on link l .

$M_{l,w}$: Primary capacity reserved on wavelength w on link l .

$R_{l,w}$: Backup capacity reserved on wavelength w on link l .

ILP Formulation I: Shared Backup Multiplexing

1. Objective: To minimize the total wavelength links with cost $C_l = 1$.

$$\min \sum_{l \in E} C_l \times W_l \quad (10.17)$$

2. Constraints on physical route variables:

$$p_{l,w}^i = \sum_{p=1}^{KW} \delta^{i,p} \epsilon_l^{i,p} \psi_w^{i,p} \quad r_{l,w}^m = \sum_{r=1}^{KW} \nu^{i,r} \epsilon_l^{i,r} \psi_w^{i,r} \quad (10.18)$$

3. Constraints on path indicators: Each request gets one primary (backup)

$$\sum_{p=1}^{KW} \delta^{i,p} = 1 \quad \sum_{r=1}^{KW} \nu^{i,r} = 1 \quad (10.19)$$

4. *Constraints on topology diversity*: Disjoint primary and backup paths.

$$\sum_{p=1}^W \delta^{i,p} = \sum_{r=W+1}^{KW} \nu^{i,r} \quad \sum_{p=W+1}^{KW} \delta^{i,p} = \sum_{r=1}^W \nu^{i,r} \quad (10.20)$$

5. *Constraints on wavelength capacity*: Wavelength capacities should not exceed.

$$M_{l,w} = \sum_i d_i \times p_{l,w}^i \quad M_{l,w} + R_{l,w} \leq C \quad (10.21)$$

6. *Constraints on fiber capacity*: $x_{l,w}$ is the sum of primary and backup paths that use wavelength w on link l . $u_{l,w} = 1$, if $x_{l,w} \geq 1$, and zero otherwise.

$$x_{l,w} = \sum_i (r_{l,w}^i + p_{l,w}^i) \quad u_{l,w} \leq x_{l,w} \quad KN(N-1)u_{l,w} \geq x_{l,w} \quad (10.22)$$

$$u_{l,w} \in \{0, 1\} \quad W_l \geq \sum_w u_{l,w} \quad W_l \leq W \quad (10.23)$$

7. *Constraints on backup multiplexing*: The capacity reserved for backup paths accounts for correlation between the corresponding primary paths to determine the maximum value. In the following $j \geq i$.

$$\begin{aligned} R_{l,w} \geq & d_i \times \nu^{i,p} \epsilon_l^{i,p} \psi_w^{i,p} \\ & + \sum_j d_n \times \nu^{j,p,i,p} \epsilon_l^{j,p} \psi_w^{j,p} \times I_{(i,\bar{p}), (j,\bar{p})} + \sum_j d_j \times \nu^{j,\bar{p},i,p} \epsilon_l^{j,\bar{p}} \psi_w^{j,\bar{p}} \times I_{(i,p), (j,\bar{p})} \\ & + \sum_j d_j \times \nu^{j,p,i,\bar{p}} \epsilon_l^{j,p} \psi_w^{j,p} \times I_{(i,\bar{p}), (j,p)} + \sum_j d_j \times \nu^{j,\bar{p},i,\bar{p}} \epsilon_l^{j,\bar{p}} \psi_w^{j,\bar{p}} \times I_{(i,p), (j,p)} \end{aligned} \quad (10.24)$$

where $\nu^{j,p,i,p}$ is a binary variable which takes the value of one when $\nu^{j,p} = 1$ and $\nu^{i,p} = 1$.

$$\nu^{j,p,i,p} \geq \nu^{j,p} + \nu^{i,p} - 1; \quad \nu^{j,p,i,p} \leq \nu^{j,p}; \quad \nu^{j,p,i,p} \leq \nu^{i,p} \quad (10.25)$$

ILP Formulation II: Dedicated Backup with MLPS

For this we change the objective function to minimize the total wavelength-links as well as total link-primary-sharing. Let C_{share}^l be the weight of s_l . Most constraints remain the same and a few change as described below.

1. Objective:

$$\min \left(\sum_{l \in E} C_l \times W_l + C_{share}^l \times s_l \right). \quad (10.26)$$

2. *Constraints on backup capacity*: Backup capacities are simply aggregated.

$$R_{l,w} = \sum_i d_i \times r_{l,w}^i \quad (10.27)$$

3. Constraints on link-primary-sharing:

$$s_l \geq \sum_i \sum_w p_{l,w}^i - 1 \quad (10.28)$$

$$s_l \leq \sum_i \sum_w p_{l,w}^i \quad (10.29)$$

10.4.4 Performance Results

The performance of grooming depends on the efficiency of *grooming fractional wavelength traffic* and the *traffic pattern* offered, and on the sequence it is offered. When traffic granularity is close to full-wavelength capacity, grooming cannot bring much improvement on wavelength utilization. Therefore, we generate traffic randomly with each request requiring 1/4 capacity of a full wavelength. Two link disjoint paths for each connection are pre-computed using the shortest-path routing algorithm. The experiments were performed on a 10-node 14-bidirectional link topology as depicted in Fig. 10.6 (right).

In one experiment with 23 requests, the formulation I (backup multiplexing) required 28 wavelength-links, while formulation II (dedicated backup) required 33 wavelength-links. In general, formulation II requires more wavelength-links than formulation I. However, in further experimentation, we also noticed that dedicated protection becomes affordable under heavy traffic scenarios as the wavelength utilization significantly improves with traffic grooming. Also formulation II is computationally less expensive and hence more practical.

10.4.5 Partial Protection

Our approaches of survivable grooming network design can be extended to the partial protection in WDM grooming networks. The grooming capability of the network makes partial protection a possible solution when the network resources are not sufficient to provide full protection for every request. It may be a preferred solution as failures are not that often and a graceful degradation of services may be acceptable when compared with no service at all even in the failure-free case. Partial protection is defined as follows.

For a request m , its requested capacity for primary or working path is given as d_m , the minimum capacity for its backup is given as b_m . The difference between partial protection and full protection is that here $0 < b_m < d_m$, while in the full protection, $b_m = d_m$. Just for the sake of completeness, when $b_m = 0$, it is called no protection for request m . The problem of partial protection is to find a primary path for request m , assigning capacity of d_m to it, and find a backup path with capacity c_m such that $b_m \leq c_m \leq d_m$. The higher the value of c_m , the better the protection request m has.

The exact ILP formulations can be modified to solve the partial protection problems in grooming networks as well. However, a direct modification makes the formulations nonlinear, because in partial protection problems, the backup capacity becomes unknown. If we reconsider the motivation of the partial protection in grooming networks, the problem might be solved differently. The main reason partial protection is adopted is that we do not have enough wavelength resource to provide full protection for each request. In other words, we may not want to exploit one extra wavelength just to provide more than the minimum capacity requirement of the backups. In this situation, the partial protection problem can be divided into two subproblems.

1. Resource minimization: given the network resource and minimum backup requirement, try to allocate each request m with primary capacity of d_m and backup capacity of b_m .
2. Protection maximization: given that all the requests are accommodated with the minimum protection requirement being satisfied, the second step is to optimally distribute the residual network capacity to provide better protection to some, if not all, of the requests.

Using the above, we propose a two-phase ILP formulation with dedicated backup reservation for the partial protection design in WDM grooming networks as follows. The details are omitted here and can be found in [14]. The experiments are performed on the same 10-node network topology. For a particular traffic matrix, a full protection requires 33 wavelength links whereas, for $P_{ratio} = 0.6$, a total of only 28 wavelength-links are required. We also noticed that many of the requests were provided with more capacity than their minimum requirements and some connection requests are fully protected. In another experiment, in the same topology with each link consisting of one fiber that carries 3 wavelengths, we experimented to evaluate blocking probability. It is assumed that random requests arrive at each node according to a Poisson process with rate λ . Each request is equally likely to be destined to any of the remaining nodes. The holding time of the requests are exponentially distributed with mean $1/\mu$. The requested capacity is uniformly distributed between a given lower- and an upper-bound, and the full wavelength capacity is chosen to be OC-48. A request is said to be accepted if and only if both a primary path and a backup path for the request can be successfully allocated.

We performed experiments with two types of traffic. In one case, the request capacity is uniformly distributed between OC-1 and OC-36. In the other case, the request capacity is uniformly distributed between OC-24 and OC-36. Fig. 10.7 presents the networking blocking performance as the node load changes. For each node load, we perform simulations in 10 rounds, with each round having 100,000 random requests. An average value is taken as the blocking probability for the given node load value. It is observed from Fig. 10.7 that as the protection ratio goes down, the network blocking performance improves. This is as expected. More importantly, there is a significant gain in blocking performance.

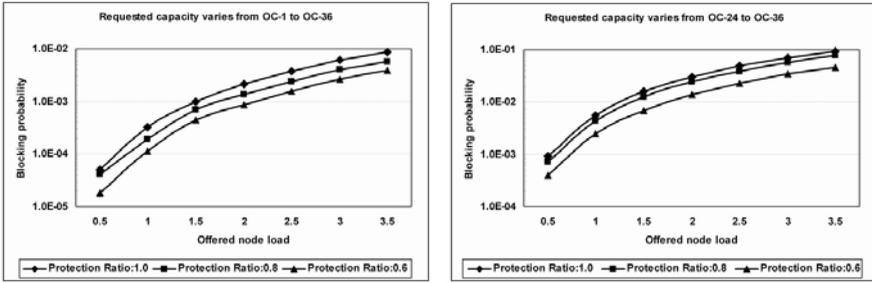


Fig. 10.7. Blocking performance for traffic (1–36) and (24–36)

10.5 Conclusion

A major issue in optical fiber network is the management of fault due to the huge amount of traffic carried by a single fiber. Capacity efficiency and recovery speed are two important aspects in designing protection mechanisms. We presented the resource planning where a single link failure is part of the design and operation process. Shared backup path protection, its variations, and p -Cycle protection are the most widely studied protection schemes in mesh network protection. Shared-path protection are believed to be capacity efficient. Studies have shown that general shared-path protection and pre-cross-connected shared-path protection use less total capacity than p -Cycle protection in low-connectivity networks, while they are comparable in high-connectivity networks.

The protection and restoration design in grooming networks is more complicated than that of the conventional WDM networks. Although shared wavelength protection is still the preferred option, our study shows that dedicated backup reservation becomes affordable and appears to be more desired when the wavelength utilization is improved by the grooming capability of the network. Furthermore, by adding a constraint to minimize the total link-primary-sharing, the number of affected working paths due to single link failure is reduced. A good practical example of designing both traffic grooming and survivability in practice is a Light-trail architecture as discussed in [18].

Dynamic establishment of restorable connections is another important issue. Due to computational complexity of the problem, developing practical and efficient heuristic approaches for survivable design remains an important research direction. When network resource is restrained and insufficient to provide 100% protection to every request, one solution is to provide partial protections. Our results show that partial protection is an effective compromise when the network resources are limited.

Acknowledgements The research is in part supported by Jerry R. Junking endowment at Iowa State University and National Science Foundation

grants CNS0323374, CNS0439748, and CNS0626741. The detailed work is reported in PhD dissertations of Drs. J. Fang [19], W. He [20], and S. Balasubramanian [21].

References

1. B. T. Doshi, S. Dravida, P. Harshavardhana, O. Hauser, and Y. Wang, "Optical Network Design and Restoration", *Bell Labs Technology Journal*, pp. 58–84, January 1999.
2. S. Ramamurthy and B. Mukherjee, "Survivable WDM mesh networks, part I: protection," *IEEE INFOCOM*, vol. 2, pp. 744–751, March 1999.
3. S. Ramamurthy and B. Mukherjee, "Survivable WDM mesh networks, part II – restoration", *IEEE ICC*, vol. 3, pp. 2023–2030, June 1999.
4. M. T. Frederick, P. Datta and A. K. Somani, "Sub-Graph Routing: A generalized fault-tolerant strategy for link failures in WDM Optical Networks," *Computer Networks* 50(2): 181–199, 2006.
5. G. Ellinas, G. Halemariam, and T. Stern, "Protection cycles in mesh WDM networks", *IEEE JSAC*, vol. 18, no. 10, pp. 1924–37, October 2000.
6. D. A. Schupke, C. G. Gruber, and A. Autenrieth, "Optimal configuration of p-cycles in WDM network", *Proceeding of IEEE ICC*, pp. 2761–2765, April 2002.
7. J. Doucette, D. He, W. D. Grover, and O. Yang, "Algorithmic approaches for efficient enumeration of candidate p-cycles and capacitated p-cycle network design," *Proceeding of DRCN 2003*, pp. 212–220, October 2003.
8. N. Jose and A. K. Somani, "Connection Rerouting/Network Reconfiguration," in Proc. of 4th DRCN 2003, Banff, Alberta, Canada, October 2003, pp. 23–30.
9. O. Gerstel and R. Ramaswami, "Optical network survivability: A service perspective", *IEEE Communication Magazine*, vol. 38, no. 3, pp. 104–113, March 2000.
10. M. Sridharan, A.K. Somani, and M.V. Salapaka, "Approaches for capacity and revenue optimization in survivable WDM networks," *Journal of High Speed Networks*, vol. 10, no. 2, pp. 109–125, August 2001.
11. G. Mohan, C. S. R. Murthy, and A. K. Somani, "Efficient algorithms for routing dependable connections in WDM optical networks," *IEEE/ACM Transactions on Networking*, vol. 9, no. 5, pp. 553–566, October 2001.
12. G. Shen and W. D. Grover, "Performance of protected working capacity envelopes based on p-cycles: Fast, simple, and scalable dynamic service provisioning of survivable services," in *Proceeding of APOC*, vol. 5626, November 2004.
13. S. Thiagarajan and A. K. Somani, "Traffic Grooming for Survivable WDM Mesh Networks," in Special Issue on Protection/Restoration Meets the Reliability Challenge, *Optical Networks Magazine*, pp. 88–98 May/June 2002.
14. J. Fang, M. J. Fang, M. Sivakumar, A. K. Somani and K. M. Sivalingam, "On Partial Protection in Groomed Optical WDM Mesh Networks," in Proc. of IEEE Dependable Systems and Networks – Dependable Computing and Communications Symposium, Japan, pp. 228–237, June 2005.

15. W. He, J. Fang and A. K. Somani, "A new p-cycle based Survivable Design for Dynamic Traffic in WDM Networks," in Proc. of IEEE Globecom 2005, November 2005.
16. A. K. Somani, "Survivability and Traffic Grooming in Optical Networks," Cambridge Press, January 2006, 436pp.
17. C. Ou, K. Zhu, H. Zhang, L. H. Shastrabuddhe, and B. Mukherjee, "Traffic grooming for survivable WDM networks – shared protection," *IEEE Journal on Selected Areas in Communications*, pp. 1367–1383, November 2003.
18. J. Fang, W. He and A. K. Somani, "Optimal light trail design in WDM optical networks," *IEEE ICC 2004*, Vol. 3, pp. 1699–1703, June 2004.
19. J. Fang, "Traffic grooming in IP over WDM optical networks," PhD thesis, Iowa State University, 2004.
20. W. He, "Survivable design in WDM mesh networks," PhD thesis, Iowa State University, 2006.
21. S. Balasubramaniam, "Design and Protection Algorithms for Path Level Aggregation of Traffic in WDM Metro Optical Networks," PhD thesis, ISU, 2007.

Traffic Grooming Under Scheduled Service

Bin Wang

11.1 Introduction

Traffic grooming offers the ability to switch low speed traffic streams into high speed bandwidth trunks so that resources are shared among multiple entities that individually need only a fraction of the resources, which optimizes the capacity utilization by allowing the resource cost to be amortized over the number of users [1, 2]. Although much of today's physical layer network infrastructure (e.g., metro networks) have been built with ring topologies using add/drop multiplexers to add/drop wavelengths and/or tributary circuits, WDM optical networks are evolving toward general mesh topologies. Indeed, much recent work has focused on grooming traffic in mesh networks. Existing work on traffic grooming has considered several types of traffic model, such as static traffic model, dynamic random traffic model, incremental traffic model, traffic matrix set model, and scheduled traffic model. In the static traffic model, all traffic demands are known in advance and do not change over time. For instance, a client company may request virtual private network capacity for connectivity among different company sites from a service provider. The objective is typically to minimize the network resources needed or cost, e.g., the amount of line terminating equipment (LTE), or to maximize the network throughput given a resource constraint, and so on. This model does not allow dynamic connection setup and tear-down. In the dynamic random traffic model, a demand is assumed to arrive at a random time and last for a random amount of time. Usually this model assumes a certain stochastic demand arrival process (e.g., Poisson process) and a demand holding time probability distribution (e.g., exponential distribution), as well as a certain spatial traffic distribution (e.g., uniform traffic). The design objective is typically to minimize the percentage of blocked traffic. Other traffic models have been considered for network planning and configuration. The work in [3] on multi-period network planning was based on an incremental traffic model and conducted network planning across several years to incrementally produce a network capable of carrying all traffic predicted up to the end of the planning

horizon. The work reported in [4] considered time-varying offered traffic in the form of a set of traffic matrices at different instants for off-line configuration so as to accommodate the time-varying traffic. The work in [5] also used a set of traffic matrices to design and dimension a WDM mesh network to groom dynamic traffic.

While these different traffic models are valid and useful in many circumstances, they are not able to capture the traffic characteristics of applications that require capacity during a specific time interval or for a certain time period. For instance, a client company may request some bandwidth from a service provider to satisfy its communication requirements at a specific time, e.g., between headquarters and production centers during office hours or between data centers during the night when backup of databases is performed. Another example is that an IP service provider may use leased static lightpaths as IP links for providing guaranteed network capacity at all times. However, during peak traffic hours (e.g., working hours of 8 am to 5 pm), the IP network needs additional links. The additional links can be realized by periodic lightpaths that are scheduled from 8 am to 5 pm daily. Other examples include many large-scale science applications (e.g., applications in high energy physics, climate simulation, astrophysics, remote control of scientific instruments, etc.) that must deliver, at scheduled time durations, hundreds of gigabits per second throughput in near future and several terabits per second within the next decade, ranging from cooperative remote visualization of massive archival data through the distribution of large amounts of simulation data, to the interactive evolution of computations through computational steering [6]. These applications require the provisioning of scheduled dedicated channels or sub-wavelength bandwidth pipes at a specific time with a certain duration. Consequently, traffic models that capture the time-limited nature of scheduled services have been proposed and studied recently. New on-demand, time-limited large-bandwidth services are beginning to be offered by operators [7]. Traffic grooming under scheduled service is essential because many applications only need low-rate connections for communication, for instance, connections between corporate sites, scheduled backup between data centers, and occupying a full wavelength for transferring a few megabits or gigabits per second of data results in very poor utilization of network resources.

Along with the scheduled traffic models, a host of issues and opportunities arise in the study of traffic grooming for network planning purpose and for service provisioning as well. In this chapter, we will survey different traffic models that capture various aspects of scheduled service, summarize some recent work conducted on traffic grooming under these models, and discuss some related issues for future research.

The rest of the chapter is organized as follows. Section 11.2 surveys traffic models for scheduled service. Recent work on traffic grooming under scheduled service is introduced in Section 11.3. Section 11.4 then discusses some related issues and future research directions of grooming under scheduled service. Section 11.5 summarizes the chapter.

11.2 Scheduled Service Traffic Model

Several forms of traffic models [8, 9, 10, 11, 12, 13] have been considered over the years to capture various aspects of scheduled service. A demand is typically characterized by a source, a destination, and a bandwidth requirement. Depending on the starting time and the holding time of the demand, four types of scheduled service are summarized as follows.

11.2.1 Demands with Known Starting Time and Known Holding Time

This type of demands has a known starting time and a known holding time. The requested bandwidth needs to be provisioned at the starting time and the connection will be torn down after the holding time ends. A specific example is the scheduled lightpath demand (SLD) model [8, 14, 15]. The objective of service provisioning is typically to find a feasible route in the network that minimizes the network resources used or maximizes the likelihood of accommodating future demands, subject to network physical constraints, such as link capacity constraint, wavelength continuity constraint, and so on. This traffic model captures many features of aforementioned applications and has been studied by many researchers [8, 15, 16, 17, 18, 19, 20, 21, 22].

11.2.2 Demands with Known Starting Time and Unknown Holding Time

This type of demands has a known starting time and its holding time is not specified. The objective of service provisioning is typically to find a feasible route in the network in advance to increase the likelihood of successful reservation. Additionally, the network resources used to accommodate the demand should be minimized. Sometimes, the holding time can be capped by a maximum time limit. This traffic model is useful for characterizing relatively long lasting future demands that expect to get as long network services as possible. This model is suitable for providing communication service for applications that are of low priority and try to take advantage of idle network resources as much as possible.

11.2.3 Demands with Unknown Starting Time and Known Holding Time

The exact starting time of this type of demands may be unknown. The holding time is, however, pre-specified. In its simple form, this type of demands can implicitly assume a starting time of current time when the demand arrives and is appropriate for characterizing dynamic demands with known holding times. In general, the demands need to be provisioned as early as possible.

This traffic model applies to applications that expect to get network services as early as possible, but can tolerate a service delay. This model is suitable for providing communication service for non-critical data backup.

11.2.4 Demands with Known Holding Time and Scheduling Time Flexibility

Many applications have a certain degree of flexibility about when exactly the scheduled demand should occur. Service providers can therefore exploit the flexibility to schedule the connections to achieve a better utilization of resources and reduce cost.

One example of this type of traffic model is the sliding scheduled traffic model proposed in [13]. A sliding scheduled demand is represented by a tuple $d = (s, t, b, \ell, r, h)$ that satisfies $r - \ell \geq h > 0$, where s and t are the source and destination, b is the number of requested capacity units, and ℓ and r are the starting time and ending time of a time window during which the demand with a holding time of h time units resides. In this model, the demand holding time h is an interval within a time window $[\ell, r]$. Rather than fixing the starting time and ending time of the demand, a scheduling time flexibility is introduced with this model. As a result, the demand is allowed to slide within a larger time window $[\ell, r]$. This model allows an application to specify a larger time window during which the demand for communication capacity is satisfied. This model gives a service provider more flexibility in provisioning the requested demand and a better opportunity to optimize the network resources since a demand is considered accommodated as long as it is provisioned within the larger time window. Given a demand d , the actual starting time of the demand is variable relative to the left boundary ℓ of its associated time window. If the demand starts at a time units after ℓ , the demand is active during $[\ell + a, \ell + a + h]$.

This model can be considered as a generalization of the three types of traffic model defined earlier. For example, a *demand with known starting time and known holding time* is a special case when the demand starting time is the starting time of the time window while the holding time is the same as the window size. A demand with known starting time and unknown holding time is also a special case in that the time window is unbounded and the holding time is the same as the time window size. The third model is yet another special case where the time window size is infinitely large.

The work of [23] considers a similar traffic model where the scheduling of periodic connections with time flexibility on a single WDM link is studied.

11.3 Traffic Grooming under Scheduled Service

The work on traffic grooming under scheduled service is more recent and relatively limited compared to work on traffic grooming in general. In this section, we introduce some recent endeavors in this area.

11.3.1 Dynamic Traffic Grooming of Subwavelength Connections with Known Duration

The work by Tornatore et al. [21] considers dynamic traffic grooming of sub-wavelength connections with known duration where the connection requests arrive one at a time with different starting time and holding time. The objective of grooming is to minimize the network resources used for accommodating each request or implicitly attempt to minimize the overall connection blocking probability.

The authors adopt an auxiliary graph (AG) $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ to expand the representation of the original network where \mathcal{V} and \mathcal{E} are the set of nodes and links, respectively. Assuming that each link has W wavelengths, the auxiliary graph \mathcal{G} is a layered graph with $W + 2$ layers where layers 1 to W denote the W wavelength layers, layer $(W + 1)$ is the lightpath layer where established lightpaths are represented, and layer $(W + 2)$ is the access layer where a traffic demand starts and terminates. When a lightpath is established, the wavelength-links used are deleted in the appropriate wavelength layers of the graph. Similarly, corresponding wavelength-links are restored in the wavelength layers when a lightpath terminates. Network resources such as transceivers, converters, wavelength-links, and lightpaths, can be modeled using proper links in the layered auxiliary graph. The cost of links is properly assigned and updated to reflect the cost of each network element, such as transceiver, converters, wavelength-link, etc. An edge is also characterized by a proper capacity. For example, a wavelength-link edge has a capacity of a wavelength whereas for a lightpath edge, its capacity is the residual capacity after accommodating traffic demands. By manipulating the cost and capacity of links, the auxiliary graph approach allows different grooming policies to be applied while taking into account various network constraints such as transceiver availability, wavelength conversion capabilities, and grooming capabilities.

Given that the connection duration information is known upon the arrival of a demand, the authors design a holding-time-aware provisioning algorithm (HTA). The connection duration information allows the residual holding time of all the existing lightpaths in the network to be evaluated. Because a lightpath usually supports multiple time-limited connections, the residual holding time of a lightpath is determined by the connection(s) with the maximum remaining holding time after which the lightpath can be terminated. With respect to the new demand, an existing lightpath, if chosen, can serve the demand if its holding time is smaller than the residual holding time of the lightpath and the remaining capacity is sufficient to accommodate the demand. If the residual holding time of the lightpath is smaller than the demand holding time, the lifetime of the lightpath may need to be prolonged in order to accommodate the demand. An efficient routing strategy is then designed based on a holding-time-aware edge cost assignment in the auxiliary graph. The idea is to assign a lower cost to an existing lightpath whose lifetime

does not need to be prolonged, and therefore encourage the routing algorithm to reuse the lightpath for grooming. Start from a network represented as an AG with uniform cost p_e for any type of link. Given a new request (s, t, b, h) where s and t are the source and destination, b is the capacity requirement, and h is the holding time, the cost of an edge e is updated using the following cost function C_e where ϵ is a small constant:

$$C_e = \begin{cases} p_e \epsilon & \text{if } e \text{ is a lightpath with residual holding time } H_r \geq h, \\ p_e \epsilon + p_e \Delta t & \text{if } e \text{ is a lightpath with residual holding time } H_r < h \\ & (\Delta t = h - H_r), \\ p_e h & \text{for all other wavelength-links in AG.} \end{cases} \quad (11.1)$$

A shortest path algorithm is run on the auxiliary layered graph with respect to the assigned link cost to groom the new demand. Once the demand is groomed, the links, link costs, and capacity will be updated accordingly.

Simulation results [21] reported by the authors indicate a significant advantage in terms of reducing blocking probability by employing the holding-time-aware grooming approach as opposed to other efficient approaches that are holding-time unaware.

11.3.2 Traffic Grooming under a Sliding Scheduled Service Model

In [12], the authors study traffic grooming under a sliding scheduled service model. No wavelength conversion is assumed in the network under consideration. And no traffic bifurcation is allowed. Given a set of sliding-scheduled traffic demands, the primary objective of traffic grooming is to minimize the total wavelength-links used while trying to meet demands' timing specifications (i.e., starting time and holding time). Because of the scheduling time flexibility offered by the sliding scheduled service model, the difficulties of the traffic grooming problem lies in the spatial and temporal constraints/flexibilities imposed/offered on the set of demands. In the spatial domain, demands are groomed in a mesh network topology and may share the same wavelength on the same link when the capacity of the wavelength allows. Grooming of demands is also subjected to the wavelength continuity constraint. In the time domain, a demand may slide within its time window and demands may overlap in time, which constrains traffic grooming when combined with the spatial constraints. However, the time-limited nature of demands offers additional opportunities for resource reuse in both time and space during traffic grooming. Obviously, reducing overlapping between demands in time helps temporal resource reuse.

A two-step approach to the traffic grooming problem is proposed and studied. First, to maximize the network resource reuse by demands in the time domain, a demand time conflict reduction algorithm is designed. Given a set of demands \mathcal{D} , the time domain conflict reduction algorithm is applied

to \mathcal{D} to find a proper placement of demand intervals in their associated time windows such that the number of demand pairs that overlap in time is minimized. This problem is solved by first constructing an interval graph \mathcal{H} based on the demands' time windows [13]. Then strong and weak edges in the graph \mathcal{H} are identified [13]. Note that strong edges reflect time conflicts of demand pairs that cannot be resolved no matter how demands are placed within their time windows whereas weak edges represent time overlapping between two demands that can be avoided if they are properly scheduled within their time windows. The proposed time conflict reduction algorithm works to remove as many weak edges as possible from the graph \mathcal{H} to obtain an edge-reduced graph \mathcal{H}' . Therefore, strong edges of \mathcal{H} are always in the resulting graph \mathcal{H}' . However, whether a weak edge of \mathcal{H} is in \mathcal{H}' depends on how demands are placed in their associated time windows.

Once demands are properly placed in their time windows to obtain a new demand set \mathcal{D}' (i.e., with demands' start time fixed), the second step is to use a time window based algorithm for traffic grooming. The design of this grooming algorithm is based on the following observations. Demands that overlap in time must be groomed with sufficient capacity in the spatial domain if the total capacity of these demands exceeds the wavelength capacity. Network resources used by one demand can be reused by another demand as long as they do not overlap in time. This motivates the authors to divide the set of demands into subsets based on the demands' starting time and ending time such that demands in different subsets are disjoint in time except for a few straddling demands. Note that it is not always possible to assign a demand to only one subset because its holding time may result in its being included in multiple subsets. Such a demand is termed as a straddling demand. Specifically, the time conflicts of demands in \mathcal{D}' are modeled using a different interval graph. Based on the interval graph representation, an efficient algorithm then divides the demands into subsets. A virtual network topology is constructed to groom each subset of demands. The resources used by demands in one subset can potentially be reused by other subsets. Grooming of a straddling demand, however, requires the same resources to be used by it across subsets.

Simulation results and comparison with a customized tabu search scheme have shown that the proposed two-step traffic grooming algorithm for sliding scheduled demands is effective in meeting demands' timing specifications as well as reducing the total network resources used [12].

11.3.3 Traffic Grooming under a Scheduled Service Model Using Light-Trails

Our recent work considers the problem of accommodating a set of subwavelength scheduled traffic demands in a shared wavelength optical network [24] (e.g., a light-trail network) with the objective of minimizing the total resources used. Specifically, consider a network $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ where \mathcal{V} is the set of nodes and \mathcal{E} is the set of links in the network. Each link has W wavelengths and a

wavelength has a capacity of C . We assume no wavelength conversion capability in the network. Let $\mathcal{D} = \{d_1, d_2, \dots, d_m\}$ be the set of scheduled traffic demands. A demand d_i is represented as $(s_i, t_i, b_i, \alpha_i, \beta_i)$ where s_i is the source node, t_i is the destination node, b_i is the requested capacity, and α_i and β_i are the starting time and the ending time of the demand. The problem is to construct a set of light-trails $\mathcal{LT} = \{lt_1, lt_2, \dots, lt_h\}$ such that all the demands in \mathcal{D} are accommodated with the objective of minimizing the total resources used, such as the number of wavelength-links used, the number of wavelengths used.

When constructing light-trails to accommodate the scheduled demands, we consider two cases: *static light-trail case* and *dynamic light-trail case*. In the static light-trail case, a light-trail will not be terminated once it is set up. In the dynamic light-trail case, a light-trail can be torn down when it is not used by any scheduled demand, so that the resources for that light-trail can be re-allocated for other light-trails. In the dynamic light-trail case, the light-trails are time limited. Integer linear programming formulations for both cases are derived. Due to the complexity of solving ILPs for large problems, heuristic iterative grooming algorithms are proposed.

Heuristic Algorithm: Static Light-Trail Case

The proposed algorithm is shown in Algorithm 1 which combines a greedy packing algorithm with evolutionary iterations. In *Heuristic_Light-Trail*, k shortest paths are generated for the source–destination pair of each demand and are included in Γ as candidate light-trails. Then light-trails are selected using *ResAlloc* (not shown) from Γ and instantiated so that every demand in the demand set is accommodated by one of the light-trails and the total cost is minimized. *ResAlloc* returns (Γ_{ins}, c) , where Γ_{ins} is the set of instantiated light-trails and c is the total cost.

The solution is improved iteratively. In each iteration, the algorithm removes a candidate light-trail lt ($lt \in \Gamma_{ins}$) from the candidate set Γ and assigns the rest of the light-trails in Γ to a new set Γ' : ($\Gamma' \leftarrow \Gamma \setminus lt$). If by using Γ' as the candidate light-trail set, algorithm *ResAlloc* produces a better solution (i.e., with a lower cost), the current best solution is replaced. The criteria of improvement is whether the cost of the current solution is reduced enough from the previous best solution by a tunable threshold θ (e.g., 0.5%). In the case where each of the light-trails in Γ_{ins} has been tried but no improvement is obtained, the algorithm breaks out from the iteration and returns the current best solution obtained. Otherwise, the iteration continues.

The idea behind *ResAlloc* is explained briefly. A candidate light-trail can carry a demand only if both the source and destination nodes of the demand are on the light-trail and the source node is on the upstream of the destination node. A demand that can be carried by a light-trail is *supportable* by that light-trail. The algorithm sorts all the demands in the demand set that are *supportable* by a candidate light-trail lt in a certain order. Different ordering schemes are possible. Studies have shown that the longest span first (LSF) is

Algorithm 1 *Heuristic_Light-Trail*(\mathcal{G}, \mathcal{D})

Require: \mathcal{G} is the network topology and \mathcal{D} is the set of scheduled demands.

```

1:  $(\Gamma_{ins}, c_0) \leftarrow (\emptyset, |\mathcal{E}| \cdot W)$ ,  $flag \leftarrow \text{TRUE}$ ;
2: Find  $k$ -shortest paths for each demand  $d_i \in \mathcal{D}$  and put them in  $\Gamma$  as the set of
   candidate light-trails;
3:  $(\Gamma_{ins}, c_0) \leftarrow \text{ResAlloc}(\mathcal{G}, \Gamma, \mathcal{D})$ ;
4: while  $flag == \text{TRUE}$  do
5:    $flag \leftarrow \text{FALSE}$ ;
6:   for all  $lt$  such that  $lt \in \Gamma_{ins}$  do
7:      $\Gamma' \leftarrow \Gamma \setminus lt$ ;
8:      $(\Gamma_{temp}, c) \leftarrow \text{ResAlloc}(\mathcal{G}, \Gamma', \mathcal{D})$ ;
9:     if  $(c_0 - c)/c_0 \geq \theta$  then
10:       $c_0 \leftarrow c$ ;
11:       $\Gamma_{ins} \leftarrow \Gamma_{temp}$ ;
12:       $flag \leftarrow \text{TRUE}$ ;
13:     end if
14:   end for
15: end while

```

most effective [25]. The span of a demand is defined as the number of hops between the source node and the destination node of the demand. The LSF scheme attempts to better utilize the space (span) of the light-trail. A demand with a span of n hops on a light-trail with a length of m ($m \geq n$) would leave other $(m - n)$ links on the light-trail idle when the demand is transmitting data. Therefore, we prefer to select the demand with a longer span when packing demands onto the light-trail so that there are fewer links being idle. If two demands have equal spans, the one with a larger requested capacity is selected first. These demands are packed on the light-trail in the same order until the wavelength capacity limit is reached or all the *supportable* demands of that light-trail have been packed.

Once all the candidate light-trails are packed with supportable demands, the utilization ratio of a light-trail lt , ρ_{lt} , is calculated as:

$$\rho_{lt} = \sum_{\forall d_i \text{ packed onto } lt} b_i \cdot (\beta_i - \alpha_i) / (C \cdot \Psi), \quad (11.2)$$

where C is the wavelength capacity and $\Psi = [|\text{Min}(\alpha_i), \text{Max}(\beta_i)|]$ and **Min**, **Max** are taken over demands packed onto light-trail lt . The algorithm then selects the candidate light-trail lt_b with the highest utilization ratio, and moves it to the set Γ_{ins} to be instantiated. All the demands that are packed onto this light-trail are removed from the demand set \mathcal{D} . Next, each remaining candidate light-trail is packed again with the demands that are left in \mathcal{D} . Based on the utilization ratio another best candidate light-trail is picked. The iteration repeats until there is no demand left in \mathcal{D} . The algorithm then assigns wavelengths to the light-trails in Γ_{ins} and calculates the total cost.

Heuristic Algorithm: Dynamic Light-Trail Case

The basic idea behind the algorithm for the dynamic case is similar to the static light-trail case. The heuristic algorithm shares the same main function of *Heuristic_Light-Trail* (Algorithm 1) with *ResAlloc* replaced by a new *ResAlloc* (not shown). A light-trail can be terminated when it is not used by any scheduled demand, so that the resources for that light-trail can be re-allocated for other light-trails. Each light-trail can therefore have several time windows during which it is active.

The question is how to determine whether a light-trail should be made active at a certain time instant. A simple answer depends on whether there is any demand being carried on the light-trail at that time instant. This approach may produce light-trails with a low capacity utilization ratio since a demand with a very small requested capacity will keep a light-trail active. Instead, we can make the light-trail inactive when it has a low capacity utilization ratio and release the resources to other light-trails to improve the resource utilization. In Fig. 11.1(a), it is shown that there are four demands packed onto the light-trail. Each demand is represented by a rectangle in the figure. The height of the rectangle indicates the requested capacity and the width indicates the time duration of the scheduled demand.

The capacity utilization of a candidate light-trail packed with demands varies over time, which can be represented by a capacity utilization curve like the one in Fig. 11.1(b). We introduce a capacity threshold σ and deem a light-trail active at a certain time instant if the capacity utilization ratio is no less than σ , or inactive otherwise (Fig. 11.1(c)). Thus a light-trail would be either active or inactive at any time instant, which can be shown by an activation

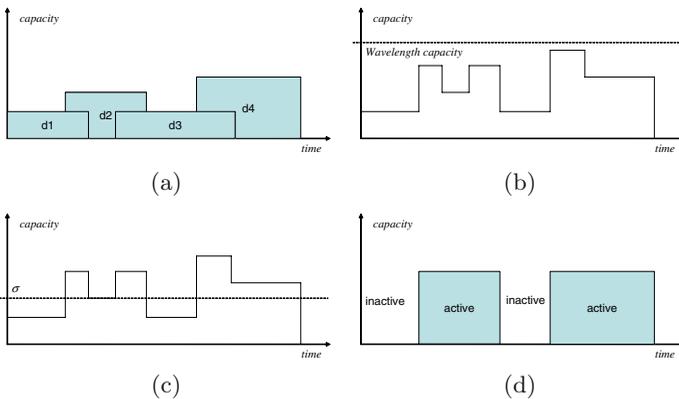


Fig. 11.1. An example of packing a time-limited light-trail with demands: (a) the demands initially packed in a light-trail; (b) the initial capacity utilization curve; (c) the capacity utilization curve with a threshold; (d) the activation curve of the light-trail

curve like the one in Fig. 11.1(d). If a demand packed in the light-trail is active during an inactive period of the light-trail, the demand will be removed from the light-trail unless the demand has the same source and destination as the light-trail, and be packed onto other potential light-trails. The demand will be accommodated by at least one light-trail because k shortest paths are generated for the source–destination pair of each demand and are included in \mathcal{L} as candidate light-trails.

The solutions to the ILP problem formulations and simulation of the heuristic algorithms show that, compared to using static light-trails, the use of dynamic light-trails for traffic grooming significantly reduces the required resources by as much as 30% for a demand set of size 100. The resource reduction becomes more significant as the size of demand sets increases and the time correlation among demands becomes weaker. The heuristic algorithms for both cases are also shown to be very time efficient. Details of simulation study can be found in [25].

11.4 Issues and Future Research Directions

The traffic grooming problem, in general, can be logically decomposed into four subproblems: (1) topology design that determines the virtual topology to be embedded in the physical topology; (2) routing that determines the route of each of the high-speed bandwidth trunks (e.g., lightpaths, light-trails, etc.) over the physical topology; (3) connection routing that routes connections over the virtual topology; and (4) wavelength assignment that allocates wavelengths to the bandwidth trunks subject to assignment and continuity constraints. In this section, we discuss related issues for grooming under scheduled service as well as future research directions.

11.4.1 Virtual Topology Design

Traditionally, designing a virtual topology on a physical network consists of deciding the lightpaths to be set up in terms of their source and destination nodes and wavelength assignment. A related problem is reconfiguring a network from one virtual topology to another. A virtual topology is designed on the basis of traffic pattern and a physical topology. Being able to reconfigure virtual topology provides adaptability, self-healing capability, and upgradability. Two types of approaches have been considered for virtual topology reconfiguration. In the cost approach, the physical network topology, the current virtual topology as well as the new virtual topology that the network must be reconfigured to are known. The goal is to minimize the cost of the reconfiguration, e.g., the cost of optical switching reprogramming, establishment of new lightpaths, and elimination of old ones. The optimization approach assumes that only the current virtual topology is given and changed

traffic pattern and/or physical topology makes reconfiguration necessary. Reconfiguration involves solving a new virtual topology design problem.

Both virtual topology design and reconfiguration have not considered the time-limited nature of scheduled service. The timing information of scheduled service demands offers another dimension for optimization as demonstrated in grooming scheduled service using light-trails (Section 11.3.3) where light-trails can be time-limited. However, significant difficulties exist in exploiting the time-limited nature of scheduled service for the betterment of traffic grooming, especially in case of dynamic traffic arrivals. The optimal virtual topology changes over time. Rather than using the best topology in every time period and incurring a possibly significant reconfiguration cost, it is necessary to consider the optimum sequence of virtual topologies that will minimize the sum of the operating and reconfiguration costs over the entire horizon. Other considerations include virtual topology design for traffic grooming [26] that (1) minimizes network resource (e.g., wavelengths, ports) usage subject to given traffic blocking requirements; and (2) maximizes performance or revenue given physical topology and resource constraints.

11.4.2 Routing

Routing is important for virtual topology design, routing of connections, and providing survivability. Not much has been done to tailor routing for scheduled service. The difficulty has to do with taking into account the connection holding time of scheduled service. The work of [27] considers the routing of holding-time aware demands under a sliding scheduled traffic model to satisfy the bandwidth, starting and holding time requirements of dynamic traffic demands. Given the network topology, current link state information, and a maximal hop count H , this Bellman-Ford flavored routing algorithm finds, for each hop count value $h, 1 \leq h \leq H$, and destination node t , all h -hop feasible path between s and t under the sliding scheduled traffic model. An h -hop feasible path between s and t is a path that satisfies both the demand's bandwidth and timing requirements. This algorithm is useful for routing scheduled connection requests over a virtual topology during traffic grooming.

11.4.3 Survivability

Survivability is a critical aspect of transport networks because of the inherent vulnerability of transmission systems. Existing work deals with survivability service provisioning in the context of scheduled lightpath service.

Kuri et al. [8, 14] address the problem of routing working and protection paths for scheduled lightpath demands in an optical transport network. The problem is formulated as a combinatorial optimization problem where the objective is to minimize the number of wavelength-links required to instantiate the lightpaths. The a priori knowledge of the SLDs' time-disjointness can be

exploited to minimize the amount of wavelength-links required to instantiate the lightpaths needed. Indeed, the same channel can be allocated to multiple lightpaths, provided that they are time-disjoint. A resource-efficient technique called backup multiplexing can be used for minimizing spare channels. In this technique, the same spare channel may serve to protect multiple lightpaths provided that two conditions do not hold simultaneously: the involved lightpaths are overlapping in time and their working paths share at least one common span. A simulated annealing (SA) based algorithm is designed to find approximate solutions to this optimization problem since finding exact solutions is computationally intractable.

The work of Tornatore et al. [20, 28, 29] exploits the connection-holding-time information to dynamically provision shared-path-protected connections. The proposed algorithm, PHOTO, takes advantage of the knowledge of the holding time of connection requests to minimize resource overbuild due to backup capacity and hence achieve resource-usage efficiency. Based on similar ideas, the authors further propose a holding-time-aware, dynamic, connection provisioning algorithm, PHOTO-GSP, to improve sharing of backup resources in segment based protection [19]. The work of Saradhi [30] considers the provisioning of fault-tolerant scheduled lightpath demands based on a two-step optimization that uses a set of pre-computed routes for working and protection paths.

In [16], we study survivable service provisioning with shared path protection under the scheduled traffic model. We consider the static version of the problem where a set of demands is given, and the setup time and tear-down time of a demand are known in advance. We study time efficient approaches to *approximating* the optimal solution to the problem. The proposed approach is based on an iterative survivable routing (ISR) scheme that utilizes a capacity provision matrix and processes demands sequentially using different demand scheduling policies. The objective is to minimize the total network resources (e.g., number of wavelength-links) used by working paths and protection paths of a given set of demands while 100% restorability is guaranteed against any single failure. The additional information on connection holding time is exploited to optimize the network resources jointly in space (i.e., backup resource sharing) and in time (i.e., taking advantage of time-disjointness amongst demands). Since a demand is considered accommodated as long as it is provisioned during its holding time, time disjoint demands (working path and protection path alike) can therefore share network resources. The proposed algorithm is evaluated against solutions obtained by integer linear programming. Simulation results indicate that the proposed ISR algorithm is very time efficient while achieving excellent performance in terms of total network resources used.

Survivable traffic grooming under scheduled service is a research area yet to be explored. A possible solution could be to route scheduled traffic demands with survivability requirements over survivable lightpaths established using the schemes described above.

11.4.4 Waveband Switching

Wavelength switching is used in current light-path networks to set up connections between node pairs. With the increase in the number of wavelengths per fiber, waveband switching (WBS) [31], wherein wavelengths are grouped into bands and switched as a single entity, has been proposed to reduce the cost and control complexity of switching nodes by minimizing the port count. WBS introduces at least one additional constraint: only the traffic carried by a fixed set of wavelengths may be grouped into a band. On the other hand, a light-trail allows the intermediate nodes along a light path to access the wavelength channel, aiming at the reduction of the total number of wavelengths. Both techniques apply traffic grooming on different levels of a WDM network. Innovative solutions that solve the waveband switching and traffic grooming problem using lightpaths or light-trails [32] in an integrated fashion can reduce the port count of multigranular optical crossconnects and the size of digital crossconnects. The solutions can further benefit from taking advantage of timing information of scheduled service. How to efficiently integrate waveband switching with traffic grooming and dynamically provision scheduled high-speed channels or end-to-end connections are open. Although a new switching unit called a waveband-label switched path is defined in GMPLS to expand the underlying provisioning capabilities of traffic grooming and wavebanding in optical networks, focused efforts are needed to investigate the underlying provisioning capabilities with traffic grooming and wavebanding as well as control plane/signaling issues related to GMPLS/ASON on advance reservation for supporting scheduled services.

11.4.5 Other Research Issues

A host of other research issues exist that are related to traffic grooming under scheduled service, such as supporting scheduled multicast service, dynamic provisioning of survivable scheduled demands, and so on. While we have considered minimizing network resource such as wavelength-links or blocking probability as our main performance metrics, other constraints (e.g., port availability, grooming capability, wavelength conversion) and objective functions (e.g., minimize add/drops, OXCs, fibers, transponders, time slots, electronic conversion) can be incorporated in the problems discussed.

11.5 Conclusions

Provisioning scheduled traffic demands has been recognized as an important class of service to be supported in future networks. Along with the scheduled traffic models proposed, a host of issues and opportunities arise in the research of traffic grooming for network planning purpose and for service provisioning as well. In this chapter, we have surveyed different traffic models that capture

various aspects of scheduled service, summarized some recent work on traffic grooming under these models, and discussed some related issues for future research.

References

1. Somani A (2006) *Survivability and Traffic Grooming in WDM Optical Networks*. Cambridge University Press
2. Balasubramanian S, Somani A (2006) On Traffic Grooming Choices for IP over WDM Networks (Invited). In: *Proceedings of BROADNETS*
3. Geary N, Antonopoulos A, Drakopoulos E, O'Reilly J (2001) Analysis of Optimization Issues In Multi-Period DWDM Network Planning. In: *Proceedings of IEEE INFOCOM*
4. Ricciato F, Salsano S, Belmonte A, Listanti M (2002) Off-line Configuration of MPLS over WDM Networks under Time-Varying Offered Traffic. In: *Proceedings of IEEE INFOCOM*
5. Srinivas N, Murthy C (2004) *Photonic Network Communications*, 7:179–191
6. Rao N, Wing W (2003) *Network Provisioning and Protocols for DOE Large-Science Applications*, Report of DOE Workshop on Ultra High-Speed Transport Protocols and Dynamic Provisioning for Large-Scale Science Applications
7. Verizon (2007) <http://www22.verizon.com/wholesale/bandwidth/>
8. Kuri J (2003) *Optimization Problems in WDM Optical Transport Networks with Scheduled Lightpath Demands*. Ph.D. Thesis, ENST Paris
9. Zheng J, Mouftah H (2001) Supporting Advance Reservation in Wavelength-Routed WDM Networks. In: *Proceedings of 10th International Conference on Computer Communications and Networks*
10. Zheng J, Mouftah H (2002) Routing and Wavelength Assignment for Advance Reservation in Wavelength-Routed WDM Optical Networks. In: *Proceedings of IEEE International Conference on Communications*
11. Zheng J, Mouftah H (2003) A Framework for Supporting Advance Reservation Service in GMPLS-Based WDM Networks. In: *Proceedings of IEEE Pacific Rim Conference on Communications, Computers and Signal Processing*
12. Wang B, Li T, Luo X, Fan Y (2004) Traffic Grooming under a Sliding Scheduled Traffic Model in WDM Optical Networks. In: *Proceedings of IEEE Workshop on Traffic Grooming in WDM Networks*
13. Wang B, Li T, Luo X, Fan Y, Xin C (2005) On Service Provisioning Under a Scheduled Traffic Model in Reconfigurable WDM Optical Networks. In: *Proceedings of BROADNETS*
14. Kuri J, Puech N, Gagnaire M (2003) Diverse Routing of Scheduled Lightpath Demands in an Optical Transport Network. In: *Proceedings of Fourth International Workshop on the Design of Reliable Communication Networks (DRCN)*
15. Kuri J, Puech N, Gagnaire M, Dotaro E, Douville R (2003) *IEEE Journal on Selected Areas in Communications*, 21:1231–1240
16. Wang B, Li T (2006) Approximating Optimal Survivable Scheduled Service Provisioning in WDM Optical Networks with Iterative Survivable Routing. In: *Proceedings of BROADNETS*

17. Li T, Wang B, Xin C, Zhang X (2005) On Survivable Service Provisioning in WDM Optical Networks under a Scheduled Traffic Model. In: Proceedings of IEEE Globecom
18. Li T and Wang B (2005) On Optimal Survivability Design in WDM Optical Networks Under a Scheduled Traffic Model. In: Proceedings of the 5th IEEE International Workshop on Design of Reliable Communication Networks (DRCN)
19. Tornatore M, Carcagni M, Ou C, Mukherjee B, Pattavina A (2006) Efficient Shared-Segment Protection Exploiting the Knowledge of Connection Holding Time. In: Proceedings IEEE Globecom
20. Tornatore M, Zhang J, Pattavina A, Mukherjee B (2005) Efficient Shared-Path Protection Exploiting the Knowledge of Connection Holding Time. In: Proceedings ONDM
21. Tornatore M, Baruffaldi A, Zhu H, Mukherjee B, Pattavina A (2007) Dynamic Traffic Grooming of Subwavelength Connections with Known Duration. In: Proceedings of OFC
22. Saradhi C, Gurusamy M (2005) Circular Arc Graph based Algorithms for Routing Scheduled Lightpath Demands in WDM Optical Networks. In: Proceedings of BROADNETS
23. Su W, Sasaki G (2003) Scheduling of Periodic Transfers with Flexibility. In: 41st Annual Allerton Conference on Communication, Control, and Computing
24. Balasubramanian S, Somani A (2006) Traffic Grooming in Statistically Shared Optical Networks. In: Proceedings of 31st IEEE Conference on Local Computer Networks
25. Luo X, Wang B (2007) Service Provisioning Under a Scheduled Traffic Model Using Light-trails in WDM Optical Networks. In: Proceedings of BROADNETS
26. Xin C, Wang B, Cao X, Li J (2006) *Journal of Lightwave Technology*, 24: 2267–2275
27. Wang B, Deshmukh (2005) *IEEE Communications Letters* 9:936–938
28. Tornatore M, Pattavina A, Mukherjee B (2005) Exploiting Connection Holding Time to Improve Resource Efficiency for Dynamic Provisioning in Shared-Path Protection. In: Proceedings OFC
29. Tornatore M, Ou C, Zhang J, Pattavina A, Mukherjee B (2005) *Journal of Lightwave Technology* 23:3138–3146
30. Saradhi C, Wei L, Gurusamy M (2004) Provisioning Fault-Tolerant Scheduled Lightpath Demands in WDM Mesh Networks. In: Proceedings of BROADNETS
31. Cao X (2004) Waveband Switching in Wavelength Division Multiplexing Networks. Ph.D. Dissertation, Department of Computer Science and Engineering, SUNY Buffalo
32. Ye Y, Woesner H, Chlamtac I (2006) *Journal of Optical Networking* 5:701–714

Dynamic Grooming Algorithms

Byrav Ramamurthy

12.1 Introduction

The emergence of Wavelength Division Multiplexing (WDM) technology provides the capability for increasing the bandwidth of Synchronous Optical Network (SONET) rings by grooming low-speed traffic streams onto different high-speed wavelength channels. Since the cost of SONET add-drop multiplexers (SADM) at each node dominates the total cost of these networks, how to assign the wavelength, groom the traffic and bypass the traffic through the intermediate nodes has received a lot of attention from researchers recently. Moreover, the traffic pattern of the optical network changes from time to time. How to develop dynamic reconfiguration algorithms for traffic grooming is an important issue.

For dynamic traffic grooming (DTG) at the network operation stage, connection requests arrive and depart dynamically. As resources have already been deployed in the network and will remain unchanged for some time, the objective of a DTG algorithm is to maximize network throughput, or minimize the blocking probability of connection requests. To achieve this objective, the grooming algorithm must provision resource-efficient routes for both lightpaths and connections.

In this chapter, we describe, in detail, dynamic traffic grooming algorithms for two cases (best-fit and full-fit) for handling reconfigurable SONET over WDM networks. For each approach, an integer linear programming model and heuristic algorithms (based on the tabu search method) are given. The results demonstrate that the tabu search heuristic can yield better solutions but has a greater running time than the greedy algorithm for the best-fit case. For the full-fit case, the tabu search heuristic yields competitive results compared with an earlier simulated annealing based method and is more stable for the dynamic case. We also highlight related work on dynamic traffic grooming algorithms in other scenarios.

SONET rings have performed well as telecommunication backbone networks for a long time [1]. Using WDM technology, multiple rings can be

supported on a single fiber ring [2]. In this architecture, each wavelength independently carries a SONET ring. Each SONET ring can further support multiple low speed streams (e.g., an OC-48 SONET ring can support 4 OC-12 or 16 OC-3 streams at the same time). At each node a WDM Add/Drop Multiplexer (WADM) adds and drops or bypasses traffic on any wavelength. At each node, there are SONET add/drop multiplexers (SADM) on each wavelength to add/drop low speed streams. So the number of SADMs per node will increase linearly with the number of wavelengths that a single fiber ring can carry. The cost of SADMs will dominate the total cost of the optical network. But in fact, it is not necessary for each node to be equipped with SADMs on each wavelength. There is a need for an SADM on a wavelength at a node only if there is traffic terminating at this node on this wavelength. For example (see Fig. 12.1), if there is a traffic stream from node A to node B through node C on wavelength λ_1 , there should be an SADM on λ_1 at node A and node B. The traffic can bypass node C without add/drop capabilities for traffic streams on λ_1 . So node C need not be equipped with an SADM on λ_1 . The problem of combining different low speed traffic streams into high-speed traffic streams in such a way that the number of SADMs is minimized is called traffic grooming. Several studies have been done on traffic grooming [3, 4, 5].

Here is an example to show that traffic grooming can reduce the number of SADMs [2]. Consider a ring network with five nodes. Each wavelength could carry two traffic streams. The traffic pattern for this example is bidirectional uniform traffic. That is, there exists the same amount of traffic in both directions for each node pair. Below gives the traffic matrix of this example.

$$T = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{bmatrix}$$

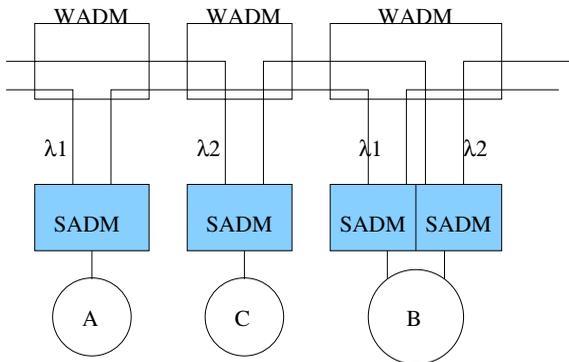


Fig. 12.1. SONET/WDM with bypass traffic at node C on λ_1

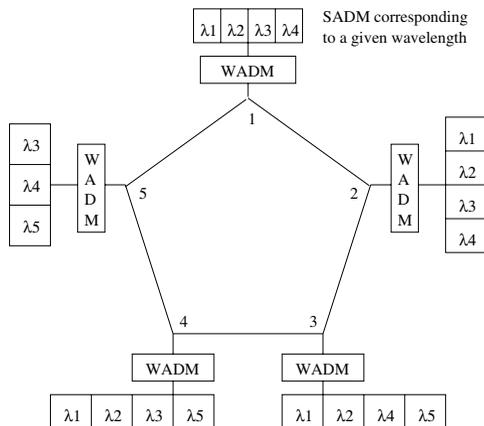


Fig. 12.2. SONET ring without traffic grooming

Figure 12.2 shows the SONET ring without traffic grooming. The traffic assignment is shown in Table 12.1. $1 \leftrightarrow 2$ in Tables 12.1 and 12.2 means that the traffic between the node 1 and 2 is groomed on the wavelength λ_1 . If there is a traffic stream from node 1 to node 3 assigned on a certain wavelength, we say that there is a virtual connection setting up from node 1 to node 3 on this wavelength. The total number of SADMs required is 19. Figure 12.3 shows the SONET ring with traffic grooming. The traffic assignment is shown in Table 12.2. The total number of SADMs is 15. In fact, if we do not know the traffic pattern and suppose there is an SADM at each node on each wavelength, the total number of SADMs is $5 \times 5 = 25$. Thus traffic grooming can reduce the number of SADMs greatly.

However, most algorithms assume the traffic matrix to be static; actually, the traffic pattern over SONET rings changes from time to time. In [6], dynamic traffic is described by a multiple set of the traffic matrices and a traffic grooming solution is proposed to meet the multiset instead of a single matrix. However, it is common that a change of traffic matrix happens after the configuration is established. In this paper, we consider the dynamic traffic grooming problem incorporating reconfiguration. That is, based on the current wavelength assignment, when the traffic pattern of the network

Table 12.1. Traffic assignment without traffic grooming

Wavelength	Traffic
λ_1	$(1 \leftrightarrow 2), (3 \leftrightarrow 4)$
λ_2	$(1 \leftrightarrow 3), (2 \leftrightarrow 4)$
λ_3	$(1 \leftrightarrow 4), (2 \leftrightarrow 5)$
λ_4	$(1 \leftrightarrow 5), (2 \leftrightarrow 3)$
λ_5	$(3 \leftrightarrow 5), (4 \leftrightarrow 5)$

Table 12.2. Traffic assignment with traffic grooming

Wavelength	Traffic
λ_1	$(1 \leftrightarrow 2), (1 \leftrightarrow 3)$
λ_2	$(1 \leftrightarrow 4), (1 \leftrightarrow 5)$
λ_3	$(2 \leftrightarrow 4), (2 \leftrightarrow 3)$
λ_4	$(4 \leftrightarrow 5), (2 \leftrightarrow 5)$
λ_5	$(3 \leftrightarrow 5), (3 \leftrightarrow 4)$

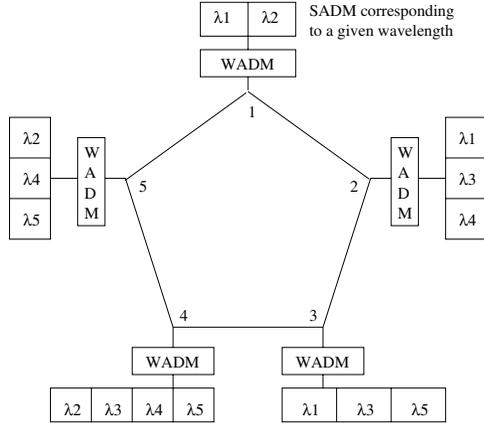


Fig. 12.3. SONET ring with traffic grooming

changes, we propose a dynamic traffic grooming algorithm to reconfigure the wavelength assignment according to the new traffic pattern without disrupting the old traffic assignment. Two cases, Best-fit and Full-fit, are studied. Two heuristic algorithms, Greedy and Tabu-search (TS-1) are presented for the best-fit approach. A two-phase algorithm based on tabu search (TS-2) is presented for the full-fit approach. The static traffic grooming problem, on which many studies have been done so far, is a special case of the *dynamic* problem (specially, the full-fit case). The simulation results illustrate that the tabu search algorithm can yield better solutions but takes more running time than greedy algorithms for the best-fit case. The tabu-search algorithm for the full-fit case is more stable compared to the earlier simulated annealing heuristic [3]. Some results for the static grooming problem are found to be better than earlier results in [3].

12.2 Problem Definition: Dynamic Traffic Grooming

In this section, integer linear programming models are proposed for dynamic traffic grooming. They are based on the static models proposed in [3]. There are two assumptions.

1. The old traffic matrix is known. By using a heuristic algorithm or Integer Linear Programming solver [3, 4], the current configuration of the network is obtained according to the old matrix. This corresponds to the initial assignment of traffic to wavelengths in the network.
2. The traffic matrix changes. The new matrix is different from the original matrix. The objective is to disrupt as few current connections as possible and fit the new traffic requests in.

The following are the notations that we will use later.

1. There are N nodes numbered $0, 1, 2, \dots, N - 1$ in the SONET ring.
2. W is the number of wavelengths in the original traffic grooming matrix.
3. The granularity of the traffic g is defined as:

$$g = \frac{\text{channel capacity}}{\text{base bandwidth rate}}$$

e.g., for a OC-48 channel carrying several OC-3 streams, the granularity $g = \frac{OC-48}{OC-3} = 16$. It is the number of circles¹ C that each wavelength can carry.

4. Original traffic matrix $T[n \times n]$. The traffic amount from node i to node j ($i, j = 1, \dots, N - 1, i \neq j$) on the ring is denoted by t_{ij} (entry of matrix T at row i and column j) and is always a multiple of the base bandwidth rate.
5. $T'[n \times n]$ is the new traffic matrix.
6. V_{ij}^{cw} represents the number of virtual connections for node pair (i, j) on wavelength w in circle c according to the original matrix. It is known in the reconfiguration problem.
7. $V'_{ij}{}^{cw}$ represents the number of new virtual connections from i to j on wavelength w of circle c according to the new traffic matrix.
8. V_{ij}^{cw+} represents the number of new virtual connections from i to j on new wavelength w^+ of circle c according to the new traffic matrix (applies to full-fit case only).
9. ADM_i^w represents the number of ADMs on node i on wavelength w for the original matrix.
10. $ADM'_i{}^w$ represents the number of additional ADMs on node i on wavelength w (applies to full-fit case only).
11. ADM_i^{w+} represents the number of ADMs on node i on new wavelength w^+ (applies to full-fit case only).

We assume that for the original traffic matrix, a solution has been found (e.g., using the algorithm in [3, 4]). So we know the number of wavelengths W and the virtual connections for each pair (i, j) . In addition, we also know the number of SADM and their positions.

Without special indication, unidirectional rings are assumed in the following problem descriptions. For bi-directional rings, constraints for both directions should be satisfied. Here are two cases to consider in the problem:

¹ The circle here refers to the circle built by the algorithm in [4].

12.2.1 Best-Fit Case

In best-fit case, we try to place as much new traffic as possible without increasing the number of SADMs.

Maximize : $\sum_i \sum_j \sum_c \sum_w V_{ij}^{\prime cw}$ (Objective function).

The objective is to maximize the traffic amount according to the new traffic demand. The following constraints are assumed.

1. Traffic constraint:

$$\sum_w \sum_c V_{ij}^{\prime cw} + \sum_w \sum_c V_{ij}^{cw} \leq t'_{ij}$$

The traffic constraint indicates that the total number of virtual connections from i to j should be less than the new demand traffic matrix of entry t'_{ij} .

2. Circle capacity constraint:

$$\sum_{(i,j)e \in (i,j)} (V_{ij}^{\prime cw} + V_{ij}^{cw}) \leq 1$$

The circle capacity constraint requires that no two connections can share a link on a circle.

3. Transmitter constraint:

$$\sum_c \sum_j V_{ij}^{\prime cw} + \sum_c \sum_j V_{ij}^{cw} \leq g \cdot ADM_i^w$$

The transmitter constraint requires that the total number of virtual connections starting at node i should be less than the transmitter capacity of SADM at this node on wavelength w .

4. Receiver constraint:

$$\sum_c \sum_i V_{ij}^{\prime cw} + \sum_c \sum_i V_{ij}^{cw} \leq g \cdot ADM_j^w$$

The receiver constraint requires that the total number of virtual connections terminating at node j should be less than the receiver capacity of SADMs at this node on wavelength w .

$V_{ij}^{cw'} \in \{0, 1, -1\}$. $ADM_j^w \in \{0, 1\}$. $V_{ij}^{cw'} = -1$ only if $V_{ij}^{cw} = 1$ and there is no connection between i and j on wavelength w of circle c for the new configuration any more.

12.2.2 Full-Fit Case

In full-fit case, we add the minimum number of SADM to satisfy all the new traffic.

Minimize : $\sum_i \sum_w ADM_i^w + \delta \sum_i \sum_{w+} ADM_i^{w+}$ (Objective function).

The objective is to fit all the traffic with the minimum number of SADMs added. δ in the objective function is the weight parameter representing the cost of adding more wavelengths. Because usually adding new wavelengths will cost more than adding more SADMs on existing wavelengths, δ is supposed to be no less than one. The following constraints are assumed.

1. Traffic constraint:

$$\sum_w \sum_c (V_{ij}^{cw} + V_{ij}^{\prime cw}) + \sum_{w+} \sum_c V_{ij}^{cw+} = t'_{ij}$$

The traffic constraint indicates that the total number of virtual connections from node i to node j should equal the traffic demand from node i to j .

2. Circle capacity constraint:

$$\sum_{(i,j) \in e \in (i,j)} (V_{ij}^{cw} + V'_{ij}{}^{cw}) \leq 1$$

$$\sum_{(i,j) \in e \in (i,j)} V_{ij}^{cw+} \leq 1$$

The circle capacity constraint requires that no two connections can share a single link on a circle.

3. Transmitter constraint:

$$\sum_c \sum_j (V_{ij}^{cw} + V'_{ij}{}^{cw}) \leq g \cdot (ADM_i^w + ADM_j'^w)$$

$$\sum_c \sum_j V_{ij}^{cw+} \leq g \cdot ADM_i^{w+}$$

The transmitter constraint requires that the total number of virtual connections should be less than the transmission capacity of the equipment at this node.

4. Receiver constraint:

$$\sum_c \sum_i (V_{ij}^{cw} + V'_{ij}{}^{cw}) \leq g \cdot (ADM_j^w + ADM_j'^w)$$

$$\sum_c \sum_i V_{ij}^{cw+} \leq g \cdot ADM_j^{w+}$$

The receiver constraint requires that the total number of virtual connections should be less than the receiving capacity of the equipment at this node.

$$V'_{ij}{}^{cw} \in \{0, 1, -1\}, V_{ij}^{cw+}, ADM_j'^w \text{ and } ADM_j^{w+} \in \{0, 1\}.$$

As we know, the integer linear programming problem is NP-complete [7]. The reconfiguration problem is described based on integer linear programming models. We expect this problem also to be intractable. In the next section, we propose heuristic approaches to solve this problem.

12.3 Heuristic Algorithms

The heuristic algorithms for dynamic grooming were developed for both the best-fit case and the full-fit case.

12.3.1 Best-Fit

The objective of the best-fit case is to include as much new traffic as possible using available capacity of the current configuration of the ring networks without increasing the number of SADMs. Two heuristic algorithms are proposed for best-fit case, greedy heuristic and Tabu search heuristic (TS-1).

Greedy Heuristic

In our greedy algorithm, the value of each entry of both the old matrix from which the current configuration was obtained and the new matrix using which

we will perform the reconfiguration is generated randomly in the range $[0, r]$. The value is uniformly distributed between 0 and r . We try to fit as much new traffic as possible without adding more SADMs. Here is a description of the algorithm.

1. Get grooming information for the original traffic matrix using an existing algorithm ([3, 4]).
For each circle, which we create for the current configuration, we should know whether the entire capacity between two nodes is occupied and whether there is an SADM at this node. We should also know on which wavelength this circle is groomed.
2. Find the difference traffic matrix.
Given the new traffic matrix, compute the difference between the old one and the new one, e.g., $D[i, j] = T'[i, j] - T[i, j]$. This is the matrix we try to groom in our algorithm with existing SADMs and traffic capacity. For some entries $D[i, j] = -m < 0$ (there exist some connections built for the old matrix that are not needed in the new matrix any more), remove the connections between (i, j) from m circles over at most m wavelengths.
3. Merge connections.
We want to keep the largest continuous gap between nodes. So, if there are two continuous connections over two circles, we merge them into a bigger one on the same circle.
4. Groom new traffic.

```

Start from the smallest hop.
From node 0 to node  $N - 1$ 
From  $s = 1$  to  $s = N / *s$  is the number of hops*/
while ( $D[i, i + s] \neq 0$ )
    Try all the circles,
    for each node pair  $(i, i + s)$  do
        if capacity is available and
        if there is SADM on both  $i$  and  $i + s$ 
        {
            Groom one unit of  $D[i, i + s]$  to this circle.
            Set capacity occupied.
             $D[i, i + s] = D[i, i + s] - 1.$ 
        }
    Endwhile
End

```

Tabu Search Heuristic (TS-1)

Tabu search is a meta-heuristic approach to solve hard optimization problems. The optimizing function is $f(x)$ subject to $x \in X$. The set X summarizes constraints on the vector of decision variables x . If x is the initial solution,

neighborhood $N(x)$ is a set obtained by going one step further from the solution x . Such a step is called a move. Each element in $N(x)$ is put into the candidate list. At the same time, a tabu list T is built to keep track of the solutions that have been visited before. Each element in the tabu list has a tabu tenure. After each move, the value of tabu tenure is decremented by one. Once the value becomes zero, the element is removed from the list. The tabu list prevents such solutions being revisited within cycles of length less than or equal to tabu tenure. A modified neighborhood $N^*(x)$ of current solution is $N(x)$ defined as $N(x) \setminus T$. Here $N^*(x)$ is a subset of $N(x)$. This kind of tabu search is short term TS. The best solution of $N^*(x)$ which is in the candidate list but not tabu (not in the tabu list) is chosen for the initial solution of next iteration. If after certain number of iterations (we call it the tabu limit) which is specified by the user, there is no improvement, the program stops. Otherwise, we continue the iteration and build a new candidate list. Here is the algorithm description of TS-1.

1. Initialization.

Compute the initial reconfiguration solution by using the greedy algorithm. Set the initial tabu tenure value, which should be defined before the program runs.

2. Build candidate list.

The neighborhood of the solution x , $N(x)$ is defined as:

$N(x) = \{x' \mid x' \text{ is a move by swapping part of the two circles of the solution } x \text{ and one part of the swap should have available capacity} \}$

Search all the circles. Compute $N(x)$ and put each element in $N(x)$ into the candidate list. Compute the additional new traffic amount that could be placed after such swap. Keep the maximum value of the increased traffic amount and its corresponding solution x .

3. Choose the best solution and continue.

For all the solutions in the candidate list, choose the move that is not tabu and can increase the traffic amount the most. The values of tabu tenure of all the elements in the tabu list are decremented by one. If some element's tenure becomes zero, remove it from the tabu list. Put the current move into the tabu list.

4. Terminate or continue with the current solution.

If after a certain number of iterations of moves, which is specified by the user, there is no improvement, the program stops. Otherwise, go back to build the candidate list for the new movement.

The tabu tenure for TS-1 is 48 and the tabu limit is 60.

12.3.2 Full-Fit Case

A two phase algorithm is developed for the full-fit case. The old traffic matrix and its solution are known. Recall that each entry of the old matrix and new

matrix is generated randomly ranging from $[0, r]$. The objective is to fit all new traffic requests. Adding more SADMs is allowed.

Algorithm Description

Here we present a two-phase algorithm for the full-fit case.

1. Use a best-fit algorithm (greedy or TS-1) to groom as much traffic as is allowed with existing SADMs.
2. If the capacity on circle c is available for the connection (i, j) , place SADMs at nodes i and j on the wavelength which circle c is groomed on and groom traffic onto this circle.
3. Use the tabu search algorithm (TS-2) to groom the remaining traffic onto the new wavelength. Place SADMs at the nodes whenever necessary.

Tabu Search Heuristic (TS-2)

The problem that was solved by TS-2 is the static traffic grooming problem, on which a lot of work has been done so far [3, 4, 5]. That is, given the remaining traffic matrix, groom that traffic onto wavelengths so that the number of SADMs is minimized. We observe that the static traffic grooming problem is a special case of the dynamic problem described in Section 12.2.2 when the old traffic matrix is empty and the new matrix is the traffic matrix that will be groomed. The following is the description of TS-2.

1. Initialization.
Use the algorithm in [4] to get an initial solution x . Set the initial value of the tabu tenure.
2. Build candidate list.
The neighborhood of the solution x , $N(x)$ is defined as:
 $N(x) = \{x' \mid x' \text{ is a move by swapping two circles from different wavelength of solution } x \text{ and the swapping will lead to no more or at most one more SADM on a wavelength}\}$.
Search all the circles. Compute $N(x)$ and place each element in $N(x)$ into the candidate list. Compute the number of SADMs that could be saved after this swap. Keep the maximum number of saved SADMs.
3. Choose the best solution and continue.
For all the solutions in the candidate list, choose the move that is not tabu and save the most SADMs. The values of tabu tenure of all the elements in the tabu list are decremented by one. If some element's tenure becomes zero, remove it from the tabu list. Place the current move into the tabu list.
4. Terminate or continue with the current solution.
If after a certain number of iterations of moves, which is specified by the user, there is no improvement, the program stops. Otherwise, go back to build the candidate list for the new movement.

The tabu limit is 170 and tabu tenure is in $\{15, 20, 24\}$.

12.4 Simulation Results and Analysis

In this section, we present the simulation results for both the best-fit case and the full-fit case.

12.4.1 Results for the Best-Fit Case

First, we define an upper bound on the number of new connections that can be groomed to the current configuration to evaluate the performance of the best-fit case algorithms. Then we give the result of both the greedy and the tabu search (TS-1) algorithms according to this upper bound. The running time comparison of both algorithms are also given.

Evaluation Factors

The results of reconfiguration algorithms depend not only on the algorithms themselves but also on the input matrix. The best-fit strategy strives to place as much traffic as it can, but there is no guarantee to fit all of the traffic in. Here we first develop an upper bound U on the number of connections that could possibly be groomed:

$$U = \sum_i \sum_j \sum_c \{V_{ij} \mid \text{there is enough capacity between } i \text{ and } j \text{ for new traffic } T'[i][j] \text{ and there is an SADM at both nodes } i \text{ and } j \text{ on circle } c\}.$$

The upper bound is computed by searching all the circles that are already built according to the old matrix to find available capacity to groom new connections. A connection can be established if there is capacity available between the terminal nodes and there are SADMs at the two nodes. The capacity of a connection is equal to the base bandwidth of one wavelength.

Although this upper bound is loose, because not all the connections available in the upper bound can be established at the same time, we can prove that no more connections can be built beyond this upper bound.

We define the load factor α by using this upper bound:

$$\alpha = \frac{\text{actual new connections established}}{\text{upper bound}} \times 100\%$$

This factor shows how much percentage of the new traffic could be groomed into the current configuration according to the upper bound.

Results of Greedy Heuristic

We randomly generate 18 matrix pairs for the old traffic matrices and the new matrices (the matrix pair is numbered as 1, 2, ..., 18 in Table 12.3 for a 5-node unidirectional ring with a granularity (g) of 3. The traffic amount generated for each node pair is evenly distributed between the range of 0 to 12 ($r = 12$). Table 12.3 provides the grooming results by using the evaluation factor α . From this table we observe that, when the number of nodes is not large and the granularity of the ring is small, the greedy algorithm can assign most of the new traffic that is possible to be groomed, according to the upper bound.

Table 12.3. Results for the best-fit greedy algorithms for 5 node ring

Matrix pair No.	1	2	3
Upper bound	15	24	24
Greedy connections	15	24	24
load factor α	100%	100%	100%
Matrix pair No.	4	5	6
Upper bound	23	20	13
Greedy connections	22	19	10
load factor α	96%	95%	77%
Matrix pair No.	7	8	9
Upper bound	10	19	8
Greedy connections	9	17	7
load factor α	90%	89%	88%
Matrix pair No.	10	11	12
Upper bound	23	12	18
Greedy connections	23	12	17
load factor α	100%	100%	94%
Matrix pair No.	13	14	15
Upper bound	22	18	33
Greedy connections	17	16	29
load factor α	82%	88%	88%
Matrix pair No.	16	17	18
Upper bound	18	16	32
Greedy connections	15	16	25
load factor α	83%	100%	78%

Results of Tabu Search Heuristic (TS-1)

Table 12.4 shows the results for the greedy algorithm and TS-1 under different numbers of nodes and different granularities for unidirectional rings ($r = 12$). For each entry, 20 matrix pairs (old matrix and new matrix) are randomly generated. The average value of α is computed. We observe that tabu search yields better results than the greedy algorithm for most cases, especially with a large number of nodes and large granularity. Table 12.5 shows the average number of new connections that can be groomed by the greedy algorithm and the tabu search algorithm using the same set of input data as Table 12.4. The tabu search heuristic gains 2% more connections than the greedy algorithm on the average.

Here we give a specific example to indicate that the tabu search heuristic can find a better solution than the greedy algorithm. In this example, there are 5 nodes in a unidirectional SONET ring. The granularity of the ring is 3. The old traffic matrix and the new matrix are generated as follows:

Table 12.4. Results for the best-fit algorithms. N is the number of nodes in a ring, α is the load factor. g is the granularity

Node number	method	g=3	g=4	g=12
N=5	greedy	92	89	79
(α in %)	tabu	92	90	80
N=6	greedy	90	89	78
(α in %)	tabu	91	90	80
N=8	greedy	90	84	76
(α in %)	tabu	91	85	77
N=12	greedy	86	83	70
(α in %)	tabu	87	84	72
N=16	greedy	85	82	67
(α in %)	tabu	86	84	68
N=20	greedy	84	80	63
(α in %)	tabu	86	82	69

$$T = \begin{bmatrix} 0 & 2 & 6 & 1 & 0 \\ 7 & 0 & 6 & 1 & 3 \\ 2 & 2 & 0 & 3 & 3 \\ 5 & 2 & 7 & 0 & 2 \\ 4 & 5 & 7 & 2 & 0 \end{bmatrix} \quad T' = \begin{bmatrix} 0 & 4 & 6 & 2 & 5 \\ 7 & 0 & 8 & 1 & 4 \\ 3 & 3 & 0 & 10 & 11 \\ 5 & 4 & 7 & 0 & 2 \\ 6 & 5 & 7 & 3 & 0 \end{bmatrix}$$

Figures 12.4 and 12.5 show the grooming results for the greedy algorithm and the tabu search algorithm, respectively. The following two matrices (R_g , R_t) are the remaining traffic matrices that cannot be groomed after employing those two algorithms. We observe from Fig. 12.5 that two connections of $t_{2,3}$ are groomed to wavelength 5 instead of wavelength 4 in the tabu search algorithm. Then two more connections of $t_{2,4}$ could be groomed to wavelength 4 which results in a gain of 2 more connections for the tabu search than the greedy heuristic.

Table 12.5. Absolute connections built by the greedy and the tabu algorithms

Node Number	method	g=3	g=4	g=12
N=5	greedy	16	17	22
	tabu	17	17	23
N=6	greedy	22	24	35
	tabu	22	25	36
N=8	greedy	39	40	54
	tabu	39	40	54
N=12	greedy	78	87	123
	tabu	79	88	126
N=16	greedy	128	142	193
	tabu	130	146	198
N=20	greedy	204	224	280
	tabu	209	229	290

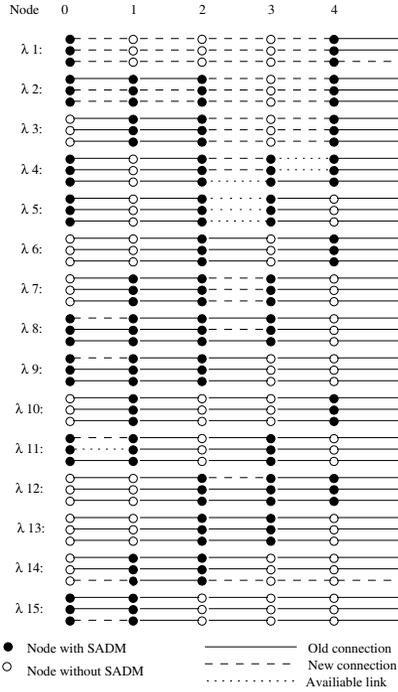


Fig. 12.4. An example of a reconfiguration problem using the greedy algorithm proposed in this work

$$R_g = \begin{bmatrix} 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 2 \\ 0 & 2 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \end{bmatrix} \quad R_t = \begin{bmatrix} 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \end{bmatrix}$$

Revenue Analysis and Running Time

While tabu search yields better results in most cases, it also takes more time to obtain the solution with the tabu heuristic than with the greedy algorithm. Figure 12.6 shows the running time of the algorithms for the unidirectional ring under different values of granularity and different numbers of nodes. It uses the same set of data that generated the result in Table 12.4. As we note from the figure, the running time increases greatly when the number of nodes in the ring increases. We compute the revenue [8] of the reconfiguration problem according to the number of connections that are built in total. Revenue is the income the network service provider earns by running the network. The following is the revenue formula for the reconfiguration problem.

C : the cost of using a single link (a single link is any node pair in $[i, i + 1]$)

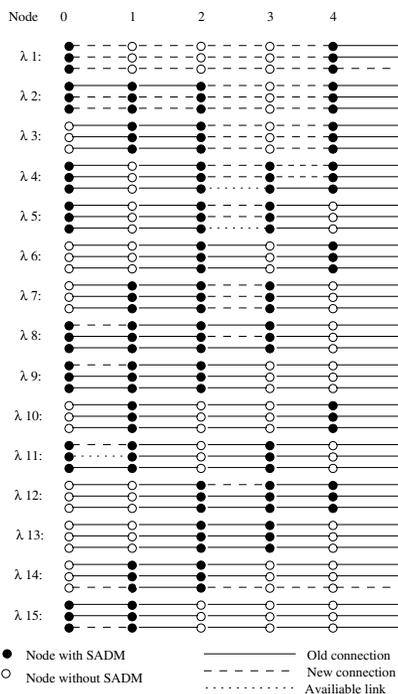


Fig. 12.5. An example of a reconfiguration problem using the tabu search heuristic (TS-1) proposed in this work

where $i \in 0..N - 1$).

$$\text{Revenue} = \sum_w \sum_c \sum_i \sum_j (V_{ij} + V'_{ij}) \times (j - i + N) \bmod N \times C \quad (12.1)$$

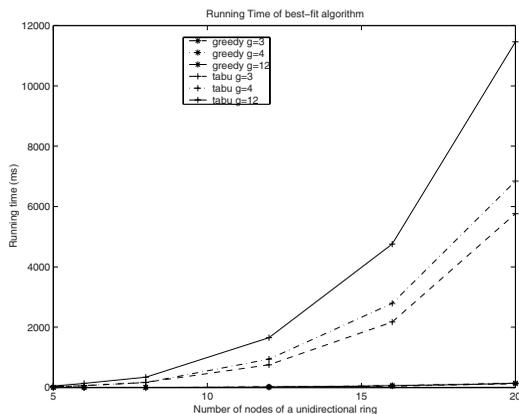


Fig. 12.6. Running time of the best-fit algorithm

From the formula, we observe that even one more connection can achieve quite a large percentage gain in the revenue if the cost of a single link is high. So, although the tabu search heuristic usually takes longer than the greedy algorithm to run, it is still worthwhile to employ tabu search when the single link cost is relatively high.

12.4.2 Results for the Full-Fit Case

We mentioned in Section 12.3.2 that the static traffic grooming problem is a special case of the dynamic problem (for the full-fit case). We run the algorithm (TS-2) for the static traffic grooming problem for a unidirectional ring under uniform traffic. The traffic amount of each entry in the matrix is one unit (equals base bandwidth). We compare the results with both the greedy algorithm [4] and the simulated annealing algorithm [3] in Table 12.6. For ILP results please refer to [3]. We observe that the tabu search algorithm obtains the same results as simulated annealing (SA) for most cases. One entry for the tabu search method is better than SA. For some cases, the number of SADMs used in tabu heuristic is a little bit more than in SA. The running time for each entry is less than 27 seconds on a 450 MHz UltraSPARC II processor based SUN Ultra-60 Workstation. In [3], the SA algorithm was run for 30 trials and best result was chosen. TS-2 does not depend on a statistical result and need not be run multiple times. We observe that it is never worse than the greedy algorithm. TS-2 is relatively stable, which is preferred in dynamic models. Because the traffic pattern changes from time to time, it is usually not feasible to run the program many times to obtain the best solution.

Table 12.6. Results for static traffic grooming

g	method	N=4	N=8	N=16	N=20
$g=3$	greedy	7	31	79	146
	SA	7	31	69	124
	TS-2	7	31	69	124
$g=4$	greedy	7	29	71	123
	SA	7	28	66	120
	TS-2	7	28	66	120
$g=16$	greedy	4	14	36	65
	SA	4	14	33	57
	TS-2	4	14	33	58
$g=48$	greedy	4	8	21	34
	SA	4	8	19	37
	TS-2	4	8	21	32
$g=64$	greedy	4	8	15	28
	SA	4	8	15	28
	TS-2	4	8	15	28

12.5 Conclusion

In this work, we introduced the dynamic traffic grooming model for reconfiguration problems. Two cases (best-fit and full-fit) are presented in an integer linear programming description. Since integer linear programming problems are NP-hard, we expect that dynamic traffic grooming problem is also intractable for both cases. For the best-fit case, two heuristic algorithms greedy and tabu search (TS-1) are proposed. An upper bound of the number of new virtual connections is developed to evaluate the performance of the algorithms. The results show that the tabu search algorithm (TS-1) proposed in this study will yield better solutions but takes more running time than the greedy algorithm for the best-fit case. For the full-fit case, a two-phase algorithm is developed. We observe that the static traffic grooming problem is a special case of the dynamic traffic grooming problem. The algorithms we proposed here (particularly the full-fit case, TS-2) can also solve the static grooming problem. Our algorithm is more stable than the simulated annealing algorithm proposed in previous work. Moreover, some of the solutions are better than those obtained in the earlier work on the static grooming problem.

As traffic grooming connects the optical layer and the electronic layer, a traffic grooming algorithm has to address not only the routing of traffic connections but also the routing and wavelength assignment (RWA) of wavelength channels. In other work, we proposed dynamic traffic grooming algorithms based on three different approaches. The first approach is an adaptive routing approach based on an auxiliary graph model [9]. The second approach is a fixed-alternate routing approach [10]. The third approach is a two-phase rerouting approach [11]. The common objective of these grooming algorithms is to maximize the network resource utilization or equivalently to minimize the connection blocking probability. We refer the interested reader to [12] and [13] and references therein for additional material on dynamic traffic grooming.

Acknowledgments The author thanks Ms. Shu Zhang and Dr. Wang Yao for their contributions to the the work presented here. This work was supported in part by the U.S. National Science Foundation (NSF).

References

1. Angela L. Chiu and Eytan H. Modiano (1998), "Reducing Electronic Multiplexing Costs in Unidirectional SONET/WDM Ring Networks via Efficient Traffic Grooming," *IEEE/IEICE Global Telecommunication Conference*, vol. 1, pp. 332–327, Sydney, Australia.
2. Eytan Modiano and Aradhana Narula-Tam (2000), "Mechanisms for providing optical bypass in WDM-based networks," *SPIE Optical Networks Magazine*.
3. Jian Wang, V. Rao Vemuri, Wonhong Cho and Biswanath Mukherjee (2000), "Improved Approaches for Cost-effective Traffic Grooming in WDM Ring networks: Non-uniform Traffic and Bidirectional Ring," *Proceedings ICC 2000*.

4. Xijun Zhang and Chunming Qiao (1998), "An Effective and Comprehensive Approach for Traffic Grooming and Wavelength Assignment in SONET/WDM Rings," *Proceedings of the SPIE-The International Society for Optical Engineering*, vol. 3531, pp. 221–232.
5. Ori Gerstel and Rajiv Ramaswami (1998), "Cost Effective Traffic Grooming in WDM rings", *Proceedings of IEEE INFOCOM '98*, vol. 1, pp. 69–77.
6. Randall Berry and Eytan Modiano (2000), "Reducing electronic multiplexing costs in SONET/WDM rings with dynamically changing traffic," *IEEE Journal on Selected Areas in Communications*, vol. 18, No. 10, pp. 1961–1971.
7. Alexander Schrijver (1986), *Theory of Linear and Integer Programming*, John Wiley & Sons, New York.
8. Murari Sridharan and Arun K. Somani (2000), "Revenue Maximization in Survivable WDM Networks," *Optical Networking and Communications, Proceeding of SPIE*, vol. 4233.
9. Wang Yao and Byrav Ramamurthy (2005), "A Novel Link Bundled Auxiliary Graph Model for Constrained Dynamic Traffic Grooming in Optical WDM Mesh Networks," *IEEE Journal of Selected Areas in Communications*, vol. 23, no. 8, pp. 1542–55.
10. Wang Yao and Byrav Ramamurthy (2004), "Dynamic traffic grooming using fixed-alternate routing in WDM mesh optical networks," in the *First Workshop on Traffic Grooming in WDM Networks, Co-located with BroadNets'2004*, San Jose, California.
11. Wang Yao and Byrav Ramamurthy (2004), "Rerouting schemes for dynamic traffic grooming in optical WDM mesh networks," in *IEEE Globecom 2004*, Dallas, Texas.
12. Wang Yao (2005), Traffic grooming in next-generation optical WDM mesh networks, Ph.D. dissertation, Department of Computer Science and Engineering, University of Nebraska-Lincoln.
13. Shu Huang and Rudra Dutta (2006), Dynamic traffic grooming: The changing role of traffic grooming, *IEEE Communications Surveys and Tutorials*, vol. 8, no. 4, pp. 32–50.
14. Peng-jun Wan, Liwu Liu and Ophir Frieder (1999), "Grooming of Arbitrary Traffic in SONET/WDM Rings," *IEEE/IEICE Global Telecommunication Conference*, vol. 1B, pp. 1012–1016.

Performance Models for Dynamic Traffic Grooming

Chunsheng Xin

13.1 Introduction

The traffic grooming problem originated from SONET and later extended into wavelength division multiplexing (WDM) optical networks. Traffic grooming in WDM optical networks is classified into *static* or *dynamic*, depending on whether the sub-wavelength client connections are given in advance, or randomly arrive/depart. Dynamic traffic grooming has been driven by several factors, e.g., the “bandwidth-on-demand” service needed by clients, and the demand to set up dynamic Label Switched Paths (LSPs) across optical networks, with the deployment of the Multi-Protocol Label Switching (MPLS) technology [1] in IP networks. While the major effort in static traffic grooming is to determine the lightpaths for aggregating the given client calls to optimize an objective, such as the number of accommodated calls [2] or the number of transponders [3], most studies on dynamic traffic grooming are to propose online algorithms to accommodate incoming sub-wavelength client connections on lightpaths that are set up on-demand in real time (e.g., see [4, 5, 6, 7, 8, 9, 10]). It is important to develop performance models for dynamic traffic grooming to evaluate the performance of grooming algorithms as well as help to study other problems in dynamic traffic grooming. A major performance metric of dynamic traffic grooming is the blocking probability of sub-wavelength connections. As such, we only discuss the analysis of blocking probability.

There is a large number of analytical studies on WDM optical networks with wavelength-granularity traffic demands, e.g., see [11, 12, 13, 14, 15, 16] and references therein. Nevertheless, these studies assumed wavelength-granularity traffic demands, and did not address sub-wavelength traffic demands. The authors in [17, 18] analyzed *WDM grooming* in time-space switched optical networks, where optical nodes are capable of sub-wavelength time-slot switching, and sub-wavelength connections can be directly set up over the time-slots in wavelength channels. In traffic grooming, optical nodes switch traffic only at the wavelength granularity, and

sub-wavelength connections have to be carried on lightpaths, e.g., see [2, 3, 4, 5, 6, 7, 8, 9, 10, 19, 20, 21]. We focus on this scenario and discuss appropriate performance models. We will utilize an approximation technique known as *reduced load approximation*, which was developed in [22, 23, 24] and has been widely adopted for analyzing general networks and optical networks.

13.2 Characterization of Dynamic Traffic Grooming

We use *calls* to refer to dynamic sub-wavelength connections from clients. The physical topology in dynamic traffic grooming includes the optical network, and client nodes connected to border optical nodes, as illustrated in Fig. 13.1. Lightpaths can be set up (and torn down) in real time between a *pair of client nodes* (node-pair). Upon receiving an incoming call request, a client node selects a sequence of lightpaths to carry this call to the destination. In this lightpath sequence, some lightpaths may not exist yet and need to be set up in real time. The grooming algorithm also tears down a lightpath if all carried calls are completed. The dynamic grooming algorithms can be generally classified into *single-hop* or *multi-hop* algorithms, depending on whether the lightpath sequence can include only one lightpath or multiple lightpaths. The single-hop grooming algorithm aggregates a call on a single lightpath to eliminate intermediate electronic processing (e.g., see [21, 25]). On the other hand, a multi-hop grooming algorithm may aggregate a call on several lightpaths, to reduce blocking probability. As a trade-off, the call needs to go through intermediate client nodes to be processed and forwarded in the electronic layer.

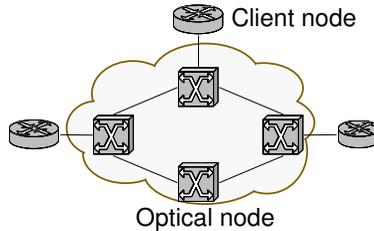


Fig. 13.1. Physical topology

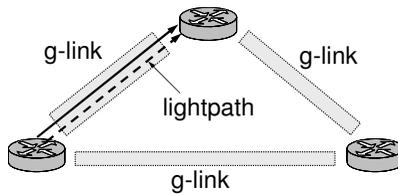


Fig. 13.2. g-link topology

The lightpaths set up in the optical network connect client nodes into a logical topology. Dynamic traffic grooming can be seen as client call routing on a dynamic logical topology, where logical links (i.e., lightpaths between node-pairs) are dynamically set up or torn down to increase or decrease capacity between a node-pair. It is difficult to analyze call blocking probabilities on a dynamic logical topology. One solution is to introduce the concept of *grooming link* (g-link) and *g-link topology* to transform a dynamic logical topology to an equivalent “static” topology. A g-link is conceived as a virtual and static “conduit” used to enclose the dynamically established lightpaths between a node-pair. The g-links between all node-pairs form a fully connected and static *g-link topology*, as illustrated in Fig. 13.2. The g-link topology together with the information on the number of lightpaths in each g-link uniquely defines a logical topology.

The single and multi-hop grooming algorithms can easily be described on the g-link topology. A single-hop grooming algorithm essentially routes a call on a single-hop path in the g-link topology, while a multi-hop grooming algorithm routes a call on a path with one or more hops.

13.3 Challenges and Solutions

To avoid confusion, a wavelength channel in a particular fiber is referred to as a *wavelink*, and a wavelink route that corresponds to a possible wavelength assignment on a fiber route is referred to as a *waveroute*. Figure 13.3 illustrates two waveroutes corresponding to two wavelength assignments on a fiber route, assuming two wavelengths per fiber and no wavelength conversion. The entire set of waveroutes (on all fiber routes) used to establish lightpaths in a g-link is referred to as the *support* of this g-link. The call loss probability between a node-pair can be calculated based on the blocking probabilities of the routing paths on the g-link topology, which are dependent on the g-link blocking probabilities. However, the number of lightpaths and the capacity of a g-link are dynamic, and it is difficult to directly calculate the g-link blocking probabilities. This problem can be solved by modeling the two operations for grooming calls on a g-link—routing calls to lightpaths and dynamically establishing lightpaths on wavelength channels for this g-link—by a single operation of routing calls directly on wavelength channels. Then the call blocking probability on a g-link becomes the joint blocking probability of the entire

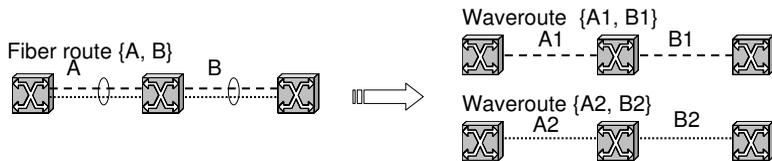


Fig. 13.3. Relationship between fiber route and waveroute

set of waveroutes supporting this g-link. However, for the single operation to have the same effect as the two operations, it has to meet two constraints.

1. Routing a call on wavelength channels must follow the routing and wavelength assignment (RWA) algorithm used to set up lightpaths, because calls are actually carried on lightpaths. Specifically, calls arriving at a g-link are routed to waveroutes that are on the fiber routes selected by the RWA algorithm to set up lightpaths in this g-link, and the waveroute selection follows the RWA scheme.
2. Calls carried on a wavelink must be on the same waveroute, to ensure that they are actually contained in a lightpath, because the traffic carried on a wavelink is actually the traffic on a lightpath that occupies the wavelink.

The second constraint can be satisfied by using a *restricted sharing* admission policy for the wavelink, which admits an incoming call if and only if the wavelink is in the empty state, or the waveroute of the incoming call is the same as the one of the carried calls and there is enough residual bandwidth.

13.4 A General Analysis Framework

Theoretically a Markov process might be used for the analysis of dynamic traffic grooming, but is far too complex. In practice, approximation techniques need to be employed. A widely used technique in network analysis called the *reduced load approximation* (see [22, 23, 24, 26]) can be adopted to analyze dynamic traffic grooming.

The objective of dynamic traffic grooming analysis is to estimate the call blocking probability for given offered loads to each node-pair. Based on the discussion in Section 13.3, we illustrate dynamic traffic grooming in Fig. 13.4. Logically, client calls are carried on the paths composed of g-links. On the other hand, client calls carried on a g-link are physically carried on waveroutes supporting this g-link, and thus carried on wavelinks. Using this hierarchy, the blocking analysis for dynamic traffic grooming is composed of three

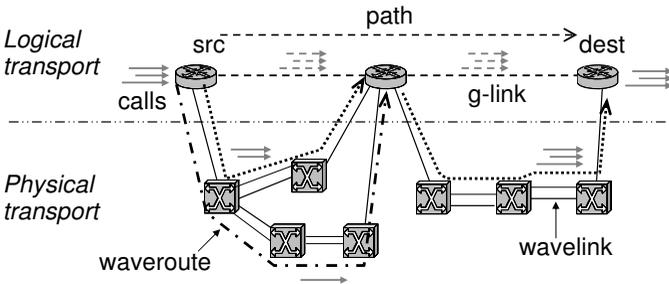


Fig. 13.4. Hierarchy in dynamic traffic grooming

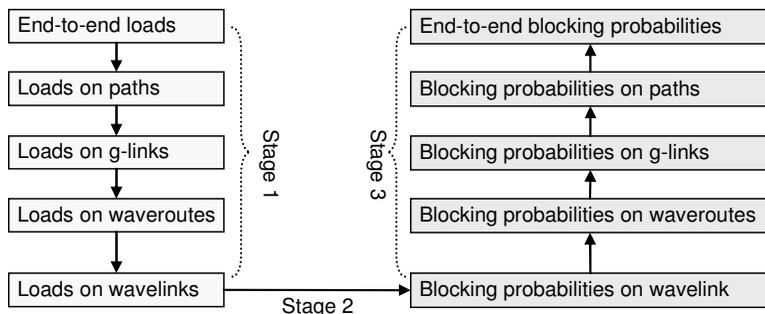


Fig. 13.5. Calculate call loss probabilities from end-to-end offered loads

stages: (1) derive wavelink offered loads from the given offered loads between node-pairs; (2) compute wavelink blocking probabilities from wavelink offered loads using a wavelink blocking model; and (3) calculate the end-to-end call loss probabilities from wavelink blocking probabilities. Figure 13.5 illustrates this procedure, where the blocking probability of a g-link is equal to the joint blocking probability of all waveroutes supporting this g-link, and the end-to-end call loss probability between a node-pair is equal to the joint blocking probability of all g-link paths used to route calls between this node-pair.

In the first stage, the end-to-end load of a node-pair is partitioned and offered to the g-link paths between this node-pair, e.g., through a sequential overflowing. The traffic load carried on a g-link is the summation of the carried loads on all paths traversing this g-link. The traffic arriving at a g-link is partitioned and offered to the waveroutes supporting this g-link, e.g., through a sequential overflowing or load sharing (each waveroute gets the same amount of traffic). Then the traffic load traversing a wavelink is simply the traffic on the waveroutes traversing this wavelink. However, this stage is a little tricky because when computing these traffic loads, the corresponding blocking probabilities are required. For instance, to calculate the traffic load carried on a specific path, we need to know both the traffic load offered to this path and the blocking probability of this path, as illustrated in Fig. 13.6. Nevertheless, our objective is to compute the blocking probabilities from these traffic loads. In other words, the blocking probabilities shown in Fig. 13.5 not only depend on the traffic loads, but also depend on themselves, which results in a set of nonlinear equations called *fixed-point equation* in the literature (e.g., see [22, 23, 24]). The fixed-point equation can be solved by a nonlinear solver and the end-to-end call loss probabilities can then be obtained.

The wavelink blocking model plays a fundamental role to obtain blocking probabilities from traffic loads. The next section introduces such a model. After this, a basic blocking model is given to illustrate the entire procedure in the analysis of traffic grooming.

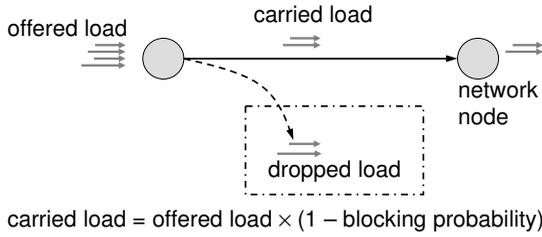


Fig. 13.6. Calculation of carried load

13.5 Wavelink Blocking Model

In traffic grooming, the calls are likely heterogeneous (with different data rates and possibly other different requirements), and hence the calls should be assumed multi-service calls, where a *service* primarily represents a data rate, but can indicate other additional requirements. The multi-service calls arriving at a wavelink are assumed following a Poisson process and having exponentially distributed holding times. In the restricted sharing admission policy¹ that is used for the wavelink, the waveroute traversed by a call is a required information for a wavelink to accept or block this call. Thus we must differentiate calls by their waveroutes. As such, we define the *class* of a call as the combination of the waveroute it traverses and its service type.

Let N denote the number of waveroutes that traverse the wavelink and $\mathcal{R}_1, \dots, \mathcal{R}_N$ denote these waveroutes. Let \mathcal{K} denote the set of call services and b_k denote the data rate of the service- k calls. Let C be the wavelink capacity. Let ρ_{nk} denote the offered load of class- $\langle n, k \rangle$ calls, i.e., the service- k calls on waveroute \mathcal{R}_n , arriving at the wavelink.

Theorem 1. Let $q(n, j)$ ($1 \leq n \leq N, 1 \leq j \leq C$) be the probability that the wavelink is occupied by calls on waveroute \mathcal{R}_n and exactly j units of bandwidth are used in the wavelink. Let $q(0)$ be the probability that the wavelink is empty (no calls).

Then $q(n, j)$ satisfies the relation

$$j \cdot q(n, j) = \sum_{k \in \mathcal{K}} b_k \cdot \rho_{nk}(j - b_k) \cdot q(n, j - b_k), \tag{13.1}$$

where $q(n, j) = 0$ for $j < 0$, $q(n, 0) = q(0)$ and

$$q(0) + \sum_{n,j} q(n, j) = 1. \tag{13.2}$$

¹ See Section 13.3.

Denote the wavelink blocking probability for class- $\langle n, k \rangle$ calls as B_{nk} . Then

$$B_{nk} = 1 - q(0) - \sum_{j=1}^{C-b_k} q(n, j), \quad (13.3)$$

Proof. See [21].

13.6 A General Blocking Model for Dynamic Traffic Grooming

Now we present a general blocking model. It can be used for general grooming algorithms that use fixed or alternate routing. The following assumptions are made:

- The optical network has no wavelength conversion.
- Link-disjoint alternate routing for both lightpaths (in the physical topology) and client calls (in the g-link topology).
- Sequential wavelength assignment, e.g., first-fit.
- Independent link blocking for wavelinks and g-links.
- A blocked lightpath/client call after trying all fiber routes/g-link paths is lost.

Based on the RWA algorithm, the waveroutes supporting each g-link are pre-computed for the analysis model. Let J denote the number of alternate g-link paths between a node-pair and U denote the number of waveroutes supporting a g-link. Let W denote the number of wavelinks in the physical topology. We use the following notations in the analysis model.

a^{sk}	Load of service- k calls offered to node-pair s .
\mathcal{L}_{sk}	Loss probability of service- k calls offered to node-pair s .
\mathcal{L}_{ave}	Average loss probability of calls (of all services) offered to the network (all node-pairs).
\mathcal{P}_{sj}	The j th path between node-pair s in the g-link topology.
\mathcal{B}^{tk}	Blocking probability of g-link t for service- k calls.
G_k^{sj}	Blocking probability of path \mathcal{P}_{sj} for service- k calls.
θ_k^{sj}	Load of service- k calls carried on path \mathcal{P}_{sj} .
C^{tk}	Load of service- k calls carried on g-link t .
A^{tk}	Load of service- k calls offered to g-link t .
\mathcal{R}_{tu}	The u th waveroute supporting g-link t .
L_k^{tu}	Blocking probability of waveroute \mathcal{R}_{tu} for service- k calls.
E_k^{tu}	Load of service- k calls carried on waveroute \mathcal{R}_{tu} (i.e., class- $\langle t, u, k \rangle$ calls).
B_w^{tuk}	Blocking probability of class- $\langle t, u, k \rangle$ calls on wavelink w .
D_w^{tuk}	Load of class- $\langle t, u, k \rangle$ calls offered to wavelink w .

13.6.1 Calculation of Traffic Loads in Stage 1

We first discuss the procedure to obtain the traffic loads on each quantity in the hierarchy discussed in Section 13.4. The calculation of the traffic loads needs the blocking probabilities. Although the actual values of these quantities are not available at this moment, they can be treated as unknowns in functions in the derivation of traffic loads.

By the alternate routing assumption, a call of node-pair s is sequentially overflowed among paths between this node-pair, $\mathcal{P}_{s,1}, \dots, \mathcal{P}_{s,J_s}$. The carried load of a path is then equal to the offered load to this path times the accepting probability, illustrated in Fig. 13.6. Thus the path \mathcal{P}_{s_j} carried load of service- k calls ($\theta_k^{s_j}$) is calculated as

$$\theta_k^{s_j} = a^{sk} \cdot \prod_{i=1}^{j-1} G_k^{s,i} \cdot (1 - G_k^{s,j}), \quad (13.4)$$

where $\prod_{i=1}^{j-1} G_k^{s,i}$ is the joint blocking probability of the first $j-1$ paths $\mathcal{P}_{s,1}, \dots, \mathcal{P}_{s,j-1}$, and thus $a^{sk} \cdot \prod_{i=1}^{j-1} G_k^{s,i}$ is the offered load to path $\mathcal{P}_{s,j}$.

The carried load on a g-link is the summation of the carried loads of all paths traversing this g-link. Let $\mathcal{Q}^t = \{(s, j)\}$ denote the indices of the paths that traverse g-link t . Then the g-link t carried load of service- k calls can be calculated as follows

$$C^{tk} = \sum_{(s,j) \in \mathcal{Q}^t} \theta_k^{s_j}. \quad (13.5)$$

The g-link t offered load (A^{tk}) can then be computed from the g-link t carried load as below (see Fig. 13.6).

$$A^{tk} = \frac{C^{tk}}{1 - \mathcal{B}^{tk}} \quad (13.6)$$

By the alternate routing and sequential wavelength assignment, a service- k call arriving at g-link t sequentially tries the waveroutes supporting g-link t , $\mathcal{R}_{t,1}, \dots, \mathcal{R}_{t,U}$. With the assumption of link-disjoint routing and no wavelength conversion in the optical network, the waveroutes $\mathcal{R}_{t,1}, \dots, \mathcal{R}_{t,U}$ are wavelink-disjoint and thus their blocking probabilities are independent. The carried load of service- k calls on waveroute \mathcal{R}_{t_u} can be computed as follows, similar to the calculation in (13.4).

$$E_k^{t_u} = A^{tk} \cdot \prod_{i=1}^{u-1} L_k^{t,i} \cdot (1 - L_k^{t,u}), \quad (13.7)$$

where $\prod_{i=1}^{u-1} L_k^{t,i}$ is the joint blocking probability of the first $u-1$ waveroutes $\mathcal{R}_{t,1}, \dots, \mathcal{R}_{t,u-1}$.

The carried load of class- $\langle t, u, k \rangle$ calls on a wavelink is the carried load of service- k calls on waveroute \mathcal{R}_{tu} , i.e., E_k^{tu} , if \mathcal{R}_{tu} traverses this wavelink. Thus the offered load of class- $\langle t, u, k \rangle$ calls to wavelink w (D_w^{tuk}) can be calculated as follows (see Fig. 13.6):

$$D_w^{tuk} = \begin{cases} \frac{E_k^{tu}}{1-B_w^{tuk}} & \text{if } w \in \mathcal{R}_{tu} \\ 0 & \text{otherwise} \end{cases}. \quad (13.8)$$

13.6.2 Calculation of Blocking Probabilities in Stage 3

After obtaining wavelink offered loads D_w^{suk} for all s, u, k , and w , wavelinks blocking probabilities B_w^{suk} can be computed using the blocking model in Section 13.5. The blocking probabilities in the third stage can then be computed.

By the independent link blocking assumption, the blocking probability of waveroutes R_{tu} for service- k calls is

$$L_k^{tu} = 1 - \prod_{w \in \mathcal{R}_{tu}} (1 - B_w^{tuk}). \quad (13.9)$$

As discussed earlier, the g-link blocking probability is equal to the joint blocking probability of all waveroutes supporting this g-link. Thus the blocking probability of g-link t (\mathcal{B}^{tk}) is as follows:

$$\mathcal{B}^{tk} = \prod_{i=1}^U L_k^{ti}. \quad (13.10)$$

By the independent link blocking, the blocking probability of path \mathcal{P}_{sj} for service- k calls is

$$G_k^{sj} = 1 - \prod_{t \in \mathcal{P}_{sj}} (1 - \mathcal{B}^{tk}). \quad (13.11)$$

By the alternate routing assumption, a call of node-pair s is sequentially overflowed among paths $\mathcal{P}_{s,1}, \dots, \mathcal{P}_{s,J}$. Thus the end-to-end call loss probability of node-pair s is

$$\mathcal{L}_{sk} = \prod_{i=1}^J G_k^{s,i}. \quad (13.12)$$

The average end-to-end call loss probability among all node-pairs is obtained by summing up dropped and offered loads on all call services and all node-pairs. Thus we have

$$\mathcal{L}_{ave} = \sum_{s,k} a^{sk} \mathcal{L}_{sk} / \sum_{s,k} a^{sk}. \quad (13.13)$$

13.6.3 Nonlinear Dependence of Blocking Probabilities

The above computation from the offered loads of node-pairs to call loss probabilities between node-pairs is not a linear procedure, since the calculation of traffic loads in Stage 1 are dependent on the blocking probabilities to be computed in Stage 3. Since all blocking probabilities in Stage 3 are calculated based on wavelink blocking probabilities, and wavelink offered loads are the final result in Stage 1, we represent the dependence of traffic loads on blocking probabilities in a function form $\mathbf{D}(\mathbf{B})$, where \mathbf{D} and \mathbf{B} represent the matrix forms of D_w^{tuk} and B_w^{tuk} . Let $\varphi(\bullet)$ denote the wavelink blocking model as a function form. Then the nonlinear dependence results in the following equation

$$\mathbf{B} = \varphi(\mathbf{D}(\mathbf{B})) = \varphi'(\mathbf{B}),$$

which is the *fixed-point equation* mentioned in Section 13.4. This equation can be solved by a nonlinear solver, e.g., through the repeated substitution illustrated in Algorithm 1. After it converges, the end-to-end call loss probabilities are obtained by (13.12)–(13.13).

We give some numerical results of the above model. A 14-node sample network illustrated in Fig. 13.7 is used. Each optical node is attached to one client node. At each client node, client calls are generated based on the Poisson arrival and exponential holding time assumptions. The wavelink capacity is assumed 100. The offered load between a node-pair is uniformly distributed

Algorithm 1 Compute call loss probabilities

1. Assign arbitrary numbers in $(0, 1)$ to B_w^{tuk} for all t, u, k, w
 2. Compute L_k^{tu} for all t, u, k by (13.9)
 3. Compute \mathcal{B}^{tk} for all k, t by (13.10)
 4. Compute G_k^{sj} for all s, j, k by (13.11)
 5. Compute θ_k^{sj} for all s, j, k by (13.4), with the given a^{sk}
 6. Compute \mathcal{C}^{tk} , and A^{tk} for all t, k by (13.5) and (13.6)
 7. Compute E_k^{tu} for all t, u, k by (13.7)
 8. Compute D_w^{tuk} for all t, u, k, w by (13.8)
 9. Compute B_w^{tuk} for all t, u, k, w by (13.1)–(13.3). If B_w^{tuk} have converged, terminate. Otherwise go to step 2
-

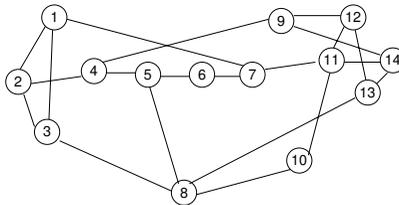


Fig. 13.7. A 14-node sample network

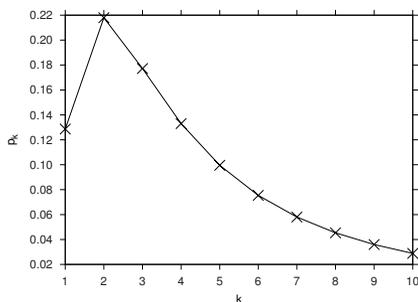


Fig. 13.8. Log normal distribution

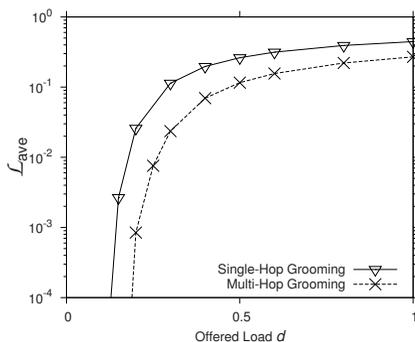


Fig. 13.9. Blocking probability

in the range $[1, 1 + d]$, where d is a simulation parameter. The number of call services is assumed 10, and the data rate of service- k ($1 \leq k \leq 10$) calls is assumed $10k$. The probability of an incoming call being a service- k call follows the log-normal distribution as illustrated in Fig. 13.8. The shortest path routing and first-fit wavelength assignment are assumed in the optical network. The number of wavelengths is assumed 8. The single-hop traffic grooming algorithm and a multi-hop grooming algorithm are used, where the latter utilizes the single-hop path and up to 4 multi-hop paths computed through setting the g-link distance as the length of the waveroute supporting the g-link. Figure 13.9 plots the computed blocking probabilities.

13.7 Extensions of the Basic Blocking Model

The blocking model in Section 13.6 can be extended to address other features and requirements in dynamic traffic grooming. We describe four examples to extend the model to address wavelength conversion, different traffic grooming policy, arbitrary alternate routing, and random traffic grooming.

13.7.1 Wavelength Conversion in Dynamic Traffic Grooming

The optical network might have wavelength conversion. With wavelength conversion, waveroutes corresponding to a fiber route overlap with each other to create *repeated wavelinks* (see Section 13.7.3), since the waveroutes are the possible wavelength assignments on this fiber route. Furthermore, the number of waveroutes corresponding to a fiber route increases exponentially as a function of the number of conversion nodes on the route. Together, they make the computation of the blocking probabilities time-consuming. To improve the efficiency, an approximation technique was proposed in [27]. A fiber route is decomposed into a sequence of *segments*,² a partial route between two consecutive conversion nodes or from the source/destination to the closest conversion node on a fiber route. As shown in Fig. 13.10, a segment is composed of multiple link-disjoint waveroutes since there is no conversion node inside a segment.

The calls offered to a fiber route arrive at each segment of the route, and are then offered to the waveroutes on each segment. The sub-wavelength calls eventually arrive at the wavelinks of these waveroutes. With this traffic grooming modeling, the blocking probability of a g-link can be calculated from the blocking probabilities of the fiber routes used to set up lightpaths for these g-links. On the other hand, the blocking probability of a fiber route can be calculated from the blocking probabilities of the segments on the route, which can be obtained from the blocking probabilities of the waveroutes on these segments. Therefore, the call loss probability between a node-pair can be computed similarly as in Section 13.6, through a hierarchy of wavelink-waveroute-segment-g-link-path. The number of waveroutes on the segments of a fiber route increases only linearly as a function of the number of conversion nodes in the fiber route. Furthermore, they are wavelink-disjoint, and also shorter than the end-to-end waveroutes on the fiber route. Therefore, the segment blocking probability can be efficiently computed.

We briefly discuss the modification to the basic blocking model in Section 13.6 for this scenario. First we redefine the following notations to make them correspond to quantities of fiber routes and then describe how to obtain them. Other notations and their computations are the same.

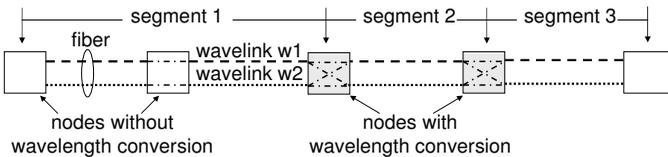


Fig. 13.10. Decomposing a fiber route into segments

² In contrast, the approach taken in Section 13.6 decomposes a fiber route into a set of end-to-end waveroutes, and the fiber route blocking probability is computed as the joint blocking probability of all waveroutes on the fiber route.

U	Number of fiber routes to set up lightpaths for a g-link.
\mathcal{R}_{tu}	The u th fiber route used to set up lightpaths on g-link t , consisting of segments.
B_w^{tuk}	Blocking probability of wavelink w for class- $\langle t, u, k \rangle$ calls (service- k calls on fiber route \mathcal{R}_{tu}).
$\mathbb{B}_{\mathbb{R}}^{tuk}$	Blocking probability of segment \mathbb{R} for class- $\langle t, u, k \rangle$ calls.
L_k^{tu}	Blocking probability of fiber route \mathcal{R}_{tu} for service- k calls.
E_k^{tu}	Carried load of service- k calls on fiber route \mathcal{R}_{tu} .
$H_{\mathbb{R}}^{tuk}$	Offered load of class- $\langle t, u, k \rangle$ calls arriving at segment \mathbb{R} .

The offered load to a g-link is offered to the fiber routes between this g-link. Furthermore, by the link-disjoint alternate routing assumption, the carried load of service- k calls on route \mathcal{R}_{tu} (i.e., class- $\langle t, u, k \rangle$ calls) is

$$E_k^{tu} = A^{tk} \cdot \prod_{i=1}^{u-1} L_k^{ti} \cdot (1 - L_k^{tu}), \quad (13.14)$$

where A^{tk} is the g-link offered load (see Section 13.6) and $\prod_{i=1}^{u-1} L_k^{ti}$ is the joint blocking probability of first $u - 1$ fiber routes.

Based on the route carried load, the offered load of class- $\langle t, u, k \rangle$ calls arriving at segment \mathbb{R} can then be calculated as

$$H_{\mathbb{R}}^{tuk} = \begin{cases} \frac{E_k^{tu}}{1 - \mathbb{B}_{\mathbb{R}}^{tuk}} & \text{if } \mathbb{R} \in \mathcal{R}_{tu} \\ 0 & \text{otherwise} \end{cases}. \quad (13.15)$$

For $\mathbb{R} \in \mathcal{R}_{tu}$, denote $|\mathbb{R}|$ as the number of waveroutes in segment \mathbb{R} (note that \mathbb{R} is composed of link-disjoint waveroutes). Let $r_1, \dots, r_{|\mathbb{R}|}$ be the waveroutes of \mathbb{R} . A call arriving at a segment tries its waveroutes sequentially, and thus the carried load of service- k calls on waveroute r_i , denoted as V_i , is

$$V_i = \prod_{\ell \in r_i} (1 - B_{\ell}^{tuk}) \cdot H_{\mathbb{R}}^{tuk} \cdot \prod_{j=1}^{i-1} \left(1 - \prod_{\ell \in r_j} (1 - B_{\ell}^{tuk}) \right),$$

where the first term is the call accepting probability on waveroute r_i , and the last term is the joint blocking probability of the first $i - 1$ waveroutes r_1, \dots, r_{i-1} on segment \mathbb{R} .

The carried load of class- $\langle t, u, k \rangle$ calls on every wavelink $w \in r_i$ is the same as the carried load of service- k calls on waveroute r_i . Thus the offered load of class- $\langle t, u, k \rangle$ calls arriving at the wavelink $w \in r_i$ ($1 \leq i \leq |\mathbb{R}|$) can be calculated as

$$D_w^{tuk} = \frac{V_i}{1 - B_w^{tuk}}.$$

After obtaining the wavelink offered loads D_w^{tuk} for all t, u, k, w , wavelink blocking probabilities B_w^{tuk} can be calculated using the wavelink blocking

model, and then other blocking probabilities can be computed based on them as follows.

By the sequential wavelength assignment assumption, a call arriving at a segment sequentially tries its waveroutes. Let $r \in \mathbb{R}$ denote a waveroute in segment \mathbb{R} . Then the call blocking probability of \mathbb{R} for class- $\langle t, u, k \rangle$ calls is calculated by the independent link blocking assumption as

$$\mathbb{D}_{\mathbb{R}}^{tuk} = \prod_{r \in \mathbb{R}} \left(1 - \prod_{w \in r} (1 - B_w^{tuk}) \right). \quad (13.16)$$

The blocking probabilities of segments on a fiber route are independent, and thus the blocking probability of fiber route \mathcal{R}_{tu} for service- k calls is

$$L_k^{tu} = 1 - \prod_{\mathbb{R} \in \mathcal{R}_{tu}} (1 - \mathbb{D}_{\mathbb{R}}^{tuk}).$$

As discussed earlier, the blocking probability of a g-link is now the joint blocking probability of the fiber routes that are used to set up lightpaths for this g-link. Thus the blocking probability of g-link t for service- k calls is

$$\mathcal{B}^{tk} = \prod_{i=1}^U L_k^{ti}. \quad (13.17)$$

Other blocking probabilities are calculated as in Section 13.6. Similarly a fixed-point equation is formed and after solving this equation, the call loss probabilities between node-pairs (\mathcal{L}_{sk}) can be obtained.

13.7.2 Traffic Grooming Policy

When routing a call onto a multi-hop g-link path, some traffic grooming algorithms may want to use only existing lightpaths on these paths, so that the source client node does not need to request to set up a new lightpath from an intermediate node to the destination node, e.g., this may not be allowed by administration policy. Moreover, requesting the setup of such a lightpath requires complicated coordination and results in more overhead. To address such constraints, the wavelink needs to differentiate a call as to whether it can request new lightpath setup or has to be carried over an existing lightpath. Thus a more sophisticated wavelink blocking model is needed to handle this situation. A call arriving at a g-link is referred to as an *active call* (with regard to this g-link) if it can request a new lightpath establishment on this g-link to accommodate it, and referred to as a *passive call* if it has to be carried on an existing lightpath on this g-link.

As illustrated in the hierarchy of dynamic traffic grooming in Fig. 13.4, a call arriving at a g-link is routed on a waveroute supporting this g-link. If a call is an active (or passive) call with regard to a g-link, it is correspondingly

referred to as an active (or passive) call on the waveroute supporting this g-link. A waveroute may carry both active and passive calls, since a g-link can carry both types of calls. Actually, the first call carried on a waveroute must be an active call that requests to establish a new lightpath on this waveroute, and later the grooming algorithm can route both active and passive calls from different node-pairs onto this waveroute, which has already become a lightpath. Besides the notations in Section 13.5, let ρ_{nk} and $\tilde{\rho}_{nk}$ denote the offered load of class- $\langle n, k \rangle$ active/passive calls arriving at the wavelink, i.e., service- k active/passive calls coming to the wavelink from waveroute \mathcal{R}_n . The following is a wavelink blocking model developed in [28] to distinguish active and passive calls.

Theorem 2. Let $\hat{\rho}_{nk}(j) = \begin{cases} \rho_{nk} & \text{if } j \leq 0 \\ \rho_{nk} + \tilde{\rho}_{nk} & \text{if } j > 0 \end{cases}$. Then $q(n, j)$ satisfies the recursive relationship

$$j \cdot q(n, j) = \sum_{k=1}^K b_k \cdot \hat{\rho}_{nk}(j - b_k) \cdot q(n, j - b_k), \quad (13.18)$$

where $q(n, j) = 0$ for $j < 0$, $q(n, 0) = q(0)$ and

$$q(0) + \sum_{n=1}^N \sum_{j=1}^C q(n, j) = 1. \quad (13.19)$$

Denote the wavelink blocking probability for class- $\langle n, k \rangle$ active/passive calls as B_{nk} and \tilde{B}_{nk} , respectively. Then

$$B_{nk} = 1 - q(0) - \sum_{j=1}^{C-b_k} q(n, j), \quad (13.20)$$

$$\tilde{B}_{nk} = 1 - \sum_{j=1}^{C-b_k} q(n, j). \quad (13.21)$$

Proof. See [28].

For the end-to-end analysis, we can define the following corresponding traffic loads and blocking probabilities for active/passive calls, similar to the notations in Section 13.6.

- C^{tk}, \tilde{C}^{tk} loads of service- k active/passive calls carried on g-link t .
- $\mathcal{B}^{tk}, \tilde{\mathcal{B}}^{tk}$ blocking probabilities of g-link t for service- k active/passive calls.
- A^{tk}, \tilde{A}^{tk} loads of service- k active/passive calls offered to g-link t .
- $L_k^{tu}, \tilde{L}_k^{tu}$ blocking probabilities of waveroute \mathcal{R}_{tu} for service- k active/passive calls.

$E_k^{tu}, \tilde{E}_k^{tu}$ loads of service- k active/passive calls carried on waveroute \mathcal{R}_{tu} (i.e., class- $\langle t, u, k \rangle$ calls).
 $B_w^{tuk}, \tilde{B}_w^{tuk}$ blocking probabilities of class- $\langle t, u, k \rangle$ active/passive calls on wavelink w .
 $D_w^{tuk}, \tilde{D}_w^{tuk}$ loads of class- $\langle t, u, k \rangle$ active/passive calls offered to wavelink w .

The traffic loads of service- k active/passive calls carried on g-link t can be calculated as follows. Let \mathcal{O}_t denote the set of node-pairs whose calls are active calls with regard to g-link t , and $\mathcal{Q}^t = \{(s, j) \mid s \in \mathcal{O}_t, \text{ and } \mathcal{P}_{s,j} \text{ traverses g-link } t\}$ denote the indices of the paths that offer active calls to g-link t . Correspondingly, we denote $\tilde{\mathcal{Q}}^t$ as the indices of the paths that offer passive calls to g-link t . Then the g-link t carried loads of service- k active and passive calls are

$$\begin{aligned} C^{tk} &= \sum_{(s,j) \in \mathcal{Q}^t} \theta_k^{sj}, \\ \tilde{C}^{tk} &= \sum_{(s,j) \in \tilde{\mathcal{Q}}^t} \theta_k^{sj}. \end{aligned}$$

The traffic loads for both active and passive calls on g-links, waveroutes, and wavelinks are calculated similarly as in Section 13.6, based on C^{tk}, \tilde{C}^{tk} . Then using the extended wavelink blocking model, we can obtain the corresponding wavelink blocking probabilities for active/passive calls. The waveroute and g-link blocking probabilities for active and passive calls are computed similarly as in Section 13.6. We define \mathcal{B}_t^{sk} as the blocking probability of g-link t for service- k calls of node-pair s , which can be computed as below.

$$\mathcal{B}_t^{sk} = \begin{cases} \mathcal{B}^{tk} & \text{if calls of node-pair } s \text{ are active calls on g-link } t \\ \tilde{\mathcal{B}}^{tk} & \text{if calls of node-pair } s \text{ are passive calls on g-link } t \end{cases}$$

The path blocking probability G_k^{sj} in (13.11) is correspondingly computed as follows

$$G_k^{sj} = 1 - \prod_{t \in \mathcal{P}_{s,j}} (1 - \mathcal{B}_t^{sk}). \quad (13.22)$$

Other blocking probabilities are calculated the same as before. The fixed-point equation formed by the nonlinear dependence of the blocking probabilities can be solved similarly and the call loss probabilities between node-pairs can then be obtained.

13.7.3 Arbitrary Alternate Routing

Both the client call routing on the g-link topology and the lightpath routing in the optical network may utilize arbitrary alternate routing instead of link-disjoint routing. The challenge in adopting arbitrary alternate routing is the *repeated g-links/wavelinks* that are present in more than one path/waveroute,

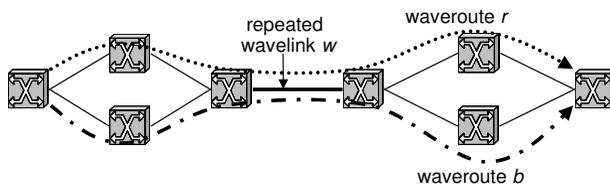


Fig. 13.11. Repeated wavelink

due to the overlapping of several paths between the same node-pair or several waveroutes supporting the same g-link, illustrated in Fig. 13.11. The repeated wavelinks/g-links make individual waveroute/path blocking probabilities not independent from other waveroutes supporting the same g-link or other paths between the same node-pair. For instance, waveroutes r and b in Fig. 13.11 have one repeated wavelink w . If a call offered to waveroute r is blocked due to wavelink w (because there is no available bandwidth), then clearly this call is also blocked on waveroute b . This makes it difficult to compute the joint waveroutes/paths blocking probability. A possible approach to address this issue is to use conditional probability law to consider the conditional joint blocking probability for each given state of the repeated wavelinks/g-links, and then use the total probability law to obtain the unconditional joint blocking probability (see [28]).

13.7.4 Random Wavelength Assignment and Random Traffic Grooming

The work in [29] discussed the analysis of random traffic grooming in optical networks with random wavelength assignment. In random traffic grooming, each client node uses one or more fiber routes to route calls to each destination. An incoming call first tries the single-hop path on a fiber route, and then randomly tries one of the two-hop g-link paths that contains only one intermediate client node on this fiber route. Figure 13.12 illustrates a fiber route from node S to D . There are three two-hop paths on this fiber route in the form of $\{S, X, D\}$, with X being A , B , or C . Two such paths are illustrated in Fig. 13.12. The analysis of this algorithm in [29] takes a similar hierarchical approach

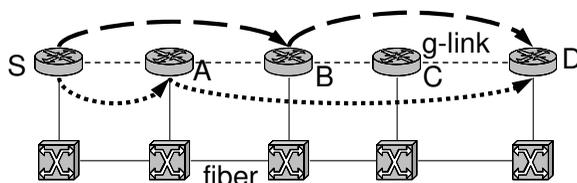


Fig. 13.12. Random traffic grooming

as in Section 13.6, with the hierarchy “single wavelength link (SWL)”–“SWL path”–lightpath–route–path, which is essentially the same hierarchy as wavelink–waveroute–g-link–path–“paths-group on a fiber route” using the terminologies of this chapter. With the random wavelength assignment, the traffic loads on the waveroutes are computed from the g-link offered load differently. The analysis model in [29] partitions the g-link offered load equally on to each waveroute. This model also divides the wavelink into *channels*, similar to time slots, and requires a client call take the same channel(s) on the entire waveroute, which is analyzed through a similar technique as analyzing the lightpath blocking in optical network without wavelength conversion (e.g., see [11]).

13.8 Conclusion

In this chapter, we first discussed the challenges for analyzing dynamic traffic grooming. We addressed these challenges by transforming traffic grooming into a problem of client call routing on a g-link topology. Then we presented an analysis framework that uses a hierarchy to help derive the client call blocking probability from the wavelink blocking probability. Next, a performance model for dynamic traffic grooming was introduced together with a wavelink blocking model that plays a key role in the performance model. At last, we demonstrated how to extend this basic model to various models for accommodating different features and requirements, e.g., wavelength conversion, traffic grooming policy, arbitrary alternate routing, and random traffic grooming.

A major research trend on performance models for dynamic traffic grooming is to enhance them to be used for studying optical network design problems with consideration of traffic grooming, e.g., topology design, routing optimization, and converter placement. Another possible future direction is to consider mixed traffic, e.g., mixed static, scheduled, and dynamic traffic, or mixed sub-wavelength, wavelength, and coarser granularity traffic, and address optical networks with multi-granularity switching, e.g., with both wavelength and waveband switching.

References

1. E. C. Rosen, A. Viswanathan, and R. Callon. (2001, Jan.) Multi-protocol label switching architecture. IETF RFC 3031. [Online]. Available: <http://www.ietf.org>
2. K. Zhu and B. Mukherjee, “Traffic grooming in an optical WDM mesh network,” *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 1, pp. 122–133, Jan. 2002.
3. J. Q. Hu and B. Leida, “Traffic grooming, routing, and wavelength assignment in optical WDM mesh networks,” in *Proc. IEEE Infocom*, pp. 495–501, 2004.

4. C. Xin, Y. Ye, S. Dixit, and C. Qiao, "An integrated lightpath provisioning approach in mesh optical networks," in *OFC Technical Digest*, pp. 547–549, 2002.
5. K. Zhu and B. Mukherjee, "Online approaches for provisioning connections of different bandwidth granularities in WDM mesh networks," in *OFC Technical Digest*, pp. 549–551, 2002.
6. H. Zhu, H. Zang, K. Zhu, and B. Mukherjee, "A novel generic graph model for traffic grooming in heterogeneous WDM mesh networks," *IEEE/ACM Transactions on Networking*, vol. 11, no. 2, pp. 285–299, Apr. 2003.
7. K. Zhu, H. Zhu, and B. Mukherjee, "Traffic engineering in multigranularity heterogeneous optical WDM mesh networks through dynamic traffic engineering," *IEEE Network*, vol. 17, no. 2, pp. 8–15, Mar./Apr. 2003.
8. W. Yao and B. Ramamurthy, "A link bundled auxiliary graph model for constrained dynamic traffic grooming in WDM mesh networks," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 8, pp. 1542–1555, Aug. 2005.
9. C. Ou, K. Zhu, H. Zang, L. Sahasrabudde, and B. Mukherjee, "Traffic grooming for survivable WDM networks—shared protection," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 9, pp. 1367–1383, Nov. 2003.
10. K. Zhu, H. Zang, and B. Mukherjee, "A comprehensive study on next-generation optical grooming switches," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 7, pp. 1173–1186, Sept. 2003.
11. A. Birman, "Computing approximate blocking probability for a class of all-optical networks," *IEEE Journal on Selected Areas in Communications*, vol. 14, no. 5, pp. 852–857, June 1996.
12. A. Sridharan and K. Sivarajan, "Blocking in all-optical networks," in *Proc. IEEE Infocom*, pp. 990–999, 2000.
13. A. Mokhtar and M. Azizoglu, "Adaptive wavelength routing in all-optical networks," *IEEE/ACM Transactions on Networking*, vol. 6, no. 2, pp. 197–206, Apr. 1998.
14. S. Subramaniam, M. Azizoglu, and A. Somani, "All-optical networks with sparse wavelength conversion," *IEEE/ACM Transactions on Networking*, vol. 4, no. 4, pp. 544–557, Aug. 1996.
15. R. Barry, "Model of blocking probability in all-optical networks with and without wavelength changer," *IEEE Journal on Selected Areas in Communications*, vol. 14, no. 5, pp. 858–867, June 1996.
16. Y. Zhu, G. Rouskas, and H. Perros, "A path decomposition approach for computing blocking probabilities in wavelength-routing networks," *IEEE/ACM Transactions on Networking*, vol. 8, no. 6, pp. 747–762, Dec. 2000.
17. R. Srinivasan and A. K. Somani, "A generalized framework for analyzing time-space switched optical networks," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 1, pp. 202–215, Jan. 2002.
18. R. Srinivasan and A. K. Somani, "Analysis of optical networks with heterogeneous grooming architectures," *IEEE/ACM Transactions on Networking*, vol. 12, no. 5, pp. 931–943, Oct. 2004.
19. V. R. Konda and T. Y. Chow, "Algorithm for traffic grooming in optical networks to minimize the number of transceivers," in *Proc. IEEE Workshop on High Performance Switching and Routing*, pp. 218–221, 2001.
20. D. Li, Z. Sun, X. Jia, and S. Makki, "Traffic grooming for minimizing wavelength usage in WDM networks," in *Proc. IEEE ICCCN*, pp. 460–465, 2002.

21. C. Xin, C. Qiao, and S. Dixit, "Traffic grooming in mesh WDM optical networks — performance analysis," *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 9, pp. 1658–1669, Nov. 2004.
22. F. P. Kelly, "Blocking probabilities in large circuit switched networks," *Advances in Applied Probability*, vol. 18, pp. 473–505, 1986.
23. S. P. Chung, A. Kashper, and K. W. Ross, "Computing approximate blocking probabilities for large loss networks with state-dependent routing," *IEEE/ACM Transactions on Networking*, vol. 1, no. 1, pp. 105–115, Feb. 1993.
24. A. Girard, *Routing and Dimensioning in Circuit-Switched Networks*. Addison-Wesley, Nov. 1990.
25. J. Bannister, J. Touch, A. Willner, and S. Suryaputra, "How many wavelengths do we really need? a study of the performance limits of packet over wavelengths," *Optical Networks Magazine*, vol. 2, pp. 1–12, Apr. 2000.
26. S. P. Chung and K. W. Ross, "Reduced load approximations for multi-rate loss networks," *IEEE Transactions on Communications*, vol. 41, no. 8, pp. 1474–1481, Aug. 1993.
27. C. Xin, J. Li, X. Cao, and B. Wang, "A route segment technique for blocking analysis of dynamic traffic grooming," in *Proc. IEEE BroadNets*, the 2nd IEEE/Create-Net International Workshop on Traffic Grooming, 2005.
28. C. Xin, "Blocking analysis of dynamic traffic grooming in mesh WDM optical networks," *IEEE/ACM Transactions on Networking*, vol. 15, no. 3, pp. 721–733, June 2007.
29. W. Yao, M. Li, and B. Ramamurthy, "Performance analysis of sparse traffic grooming in WDM mesh networks," in *Proc. IEEE ICC*, pp. 1766–1770, 2005.

Multipoint Traffic Grooming

Ahmed E. Kamal, Ahmad N. Khalil

14.1 Introduction

Several of the new and emerging applications using high-performance networks use one of the multipoint service modes. Under multipoint service, and in the same session, there may be multiple sources, multiple destinations, or both. Multipoint communication can take one of the following forms [1]:

1. *One-to-Many or Multicast*: This type of service is very well known, and several applications belong to this class. These include document distribution, on-demand video distribution, network news distribution, and file distribution and caching.
2. *Many-to-One*: This type of service corresponds to data delivered from a group of users to a single destination. Applications of this type include resource discovery, data collection at a central location, auctions, group polling, and accounting.
3. *Many-to-Many*: In this type of service, several users interact together. Applications requiring this type of traffic include the combination of one-to-many and many-to-one interactions cited above between a speaker and a group of users. They also include multimedia conferencing, synchronized resources, distance learning, distributed simulations, and collaborative processing.

In this chapter, we address the problem of grooming multipoint sessions. Since this area of research is very young, very few results are available. We therefore only consider the grooming of traffic generated in multicast and many-to-one sessions.

This research was supported in part by grant CNS-0626741 from the National Science Foundation.

14.2 Multicast Traffic Grooming

Providing multicast service on optical networks was traditionally considered in two domains, namely, optical multicasting on wavelength routing networks, where a session uses a full wavelength capacity [2, 3], and multicasting on passive star optical couplers in broadcast and select networks [4]. In this section, we consider multicasting on second generation optical networks, i.e., wavelength routed networks, where signals from different multicast sessions have to be groomed. To support this type of service, traffic may have to be duplicated. Electronic equipment functionalities are therefore different, and network design, as well as session provisioning must take this into account. Techniques developed for multicasting in optical networks, which use pure optical multicasting are not applicable to multicast traffic grooming. This is due to at least two reasons: (1) since there is no traffic grooming involved in optical multicasting, and (2) since traffic duplication is implemented in the optical domain using optical splitters, which applies to all traffic on the wavelength channel, even if no branching is required by several sessions. We illustrate the concept of traffic grooming, and the importance of optimally provisioning multicast sessions, using the examples in Fig. 14.1 in which three multicast sessions have to be provisioned over a bidirectional ring with six nodes, A through F. The multicast sessions are given by:

Session 1: Source = A; Destination = {B, C}; Traffic demand = 1 unit;

Session 2: Source = B; Destination = {C}; Traffic demand = 2 unit;

Session 3: Source = A; Destination = {F}; Traffic demand = 1 unit;

Each wavelength on the ring can accommodate two traffic units. The two Figures illustrate two different ways of accommodating the sessions, which result in different levels of resource requirements. For example, in Fig. 14.1.(a), 7 units of Line Terminating Equipment (LTE) and two wavelengths are used, while the routing in Fig. 14.1.(b) costs only 6 LTEs and one wavelength.

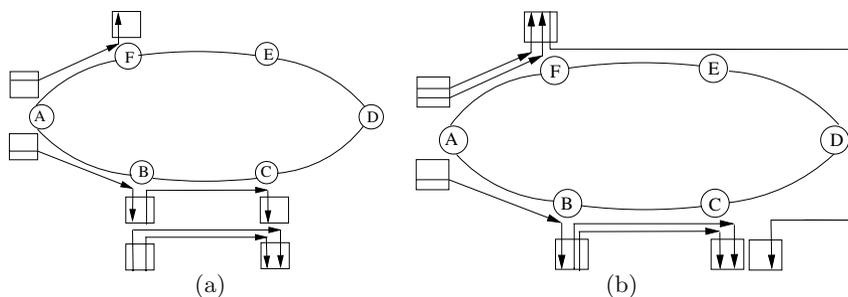


Fig. 14.1. Two different ways of multicast traffic grooming on a ring network

14.2.1 Enabling Technology

When multicast traffic grooming is involved, it may happen that at a node in the network, some of the tributaries aggregated on a certain wavelength need to be duplicated, while others need not be duplicated. In this case, it is natural to use an approach in which the optical signal is terminated at an LTE, and the tributaries are accessed. Tributaries that need to be copied are then duplicated in the electronic domain. The LTE shown in Fig. 14.2 performs this operation.

The implementation of the LTEs in Fig. 14.2 is more suitable when the LTEs are SONET ADMs. However, when using switching nodes supporting Data over SONET (DoS) in Next Generation SONET (NGS) networks [6], e.g., the DoS node shown in Fig. 14.3, two issues will be different. First, an integrated STM/packet switching architecture is used, hence obviating the need for digital cross connects. Second, data duplication may not be needed since data is encapsulated into Generic Framing Procedure (GFP) frames, and frames may be transmitted on different ports, provided that the overhead is adjusted to reflect the new destination. Reference [7] introduced a multicast extension header for GFP.

In references [5, 8], it was argued that implementing some of the traffic duplication in the optical domain might be less expensive, since the cost of (passive) optical splitters is considerably less than the cost of electronic LTEs. This is especially true if the duplicated traffic continues to use the same wavelength on different OXC output ports, and if the LTEs are only used for traffic duplication, and not for traffic dropping at a destination. However, electronic devices are still needed if traffic needs to be added to wavelength channels. Moreover, optical splitting can result in overusage of wavelength channels. Therefore, the authors introduced the node architecture shown in

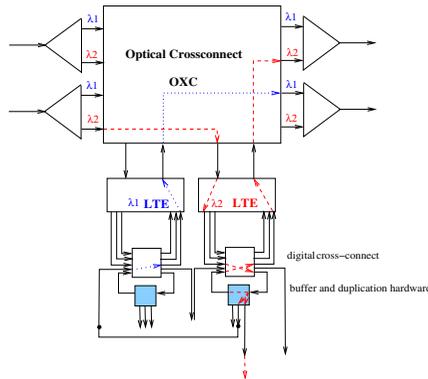


Fig. 14.2. Multicast traffic duplication in the electronic domain

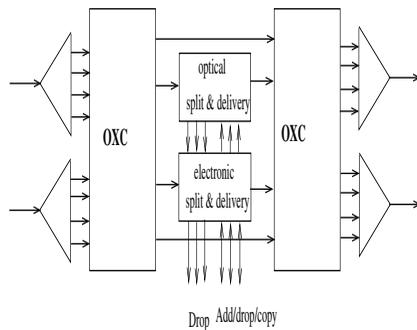


Fig. 14.3. Node architecture employing optical and electronic duplication proposed in [5]

Fig. 14.3 which implements both electronic and optical duplication, based on need and cost. Such nodes are known as translucent nodes.

14.2.2 Network Model

We consider a network with N nodes, where the network graph can assume any arbitrary mesh topology. Each edge in the graph corresponds to a pair of fibers between a pair of nodes, which are used for transmission in two different directions. Each fiber has W wavelength channels. The capacity of each channel is expressed in g , the grooming factor, which is the number of basic units of traffic which can be carried on a channel. For example, if the transmission rate on a wavelength channel is OC-48, or 2.4 Gbps, and the basic transmission unit is OC-3, or 150 Mbps, then $g = 16$. There are K multicast sessions in the network. Each session, a , where $1 \leq a \leq K$, is identified by the tuple $\langle s_a, D_a, m_a \rangle$. s_a is the source of the session, D_a is the set of destinations of the session, and m_a is the number of traffic units to be delivered from s_a to all destinations $d \in D_a$.

14.2.3 Static Multicast Traffic Grooming: Exact Approaches

Integer Linear Program (ILP)

We first consider that all traffic duplication is carried out in the electronic domain, using the node architecture shown in Fig. 14.2. An exact approach for the static multicast traffic grooming problem was presented in [7] assuming this node architecture. We present a simplified version of this approach, in the form of an Integer Linear Program (ILP), which minimizes the number of electronic equipment, LTEs. Table 14.1 presents the basic parameters needed for ILP formulation. The ILP can be formulated as follows:

Table 14.1. Multicast traffic grooming ILP basic parameters

Q	A very large integer number, (in our case $Q \geq N^2 - N$)
F_{mn}	a binary variable that indicates the presence of an edge between nodes m and n , $1 \leq m, n \leq N$: equals 1 if and only if nodes m and n are adjacent
$F_{mn}^{ij,w}$	number of lightpaths between node pair (i, j) routed on fiber (m, n) on wavelength w
T_n	number of LTEs at node n
L_{ij}^w	number of lightpaths from node i to node j on wavelength w
L_{ij}	number of lightpaths from node i to node $j = \sum_w L_{ij}^w$
$Z_{ij}^{a,d}$	a binary indicator: equals 1 if and only if session a , destined to d , is employing a lightpath from i to j as an intermediate virtual link
M_{ij}^a	a binary indicator: equals 1 if and only if $\exists d \in D_a$, such that $Z_{ij}^{a,d} = 1$. This means that session a is using lightpath (i, j) to reach at least one destination.

Objective function:

$$\text{Minimize : } \sum_{1 \leq n \leq N} T_n \quad (14.1)$$

Subject to:

- *Number of LTEs:*

$$T_i \geq \sum_w \sum_{j, j \neq i} L_{ij}^w \quad \forall i \quad (14.2)$$

$$T_i \geq \sum_w \sum_{j, j \neq i} L_{ji}^w \quad \forall i \quad (14.3)$$

- *Lightpath level constraints:*

$$\sum_{m, \text{for } F_{mi}=1} F_{mi}^{ij,w} = 0 \quad \text{and} \quad \sum_{n, \text{for } F_{in}=1} F_{jn}^{ij,w} = 0 \quad \forall i, j, w \quad (14.4)$$

$$\sum_{m, \text{for } F_{mx}=1} F_{mx}^{ij,w} = \sum_{n, \text{for } F_{xn}=1} F_{xn}^{ij,w} \quad \forall w, i, j, x; x \neq i, j \quad (14.5)$$

$$\sum_{m, \text{for } F_{mj}=1} F_{mj}^{ij,w} = L_{ij}^w \quad \text{and} \quad \sum_{n, \text{for } F_{in}=1} F_{in}^{ij,w} = L_{ij}^w \quad \forall i, j, w \quad (14.6)$$

$$\sum_i \sum_{j, \text{for } F_{mn}=1} F_{mn}^{ij,w} \leq 1 \quad \forall m, n, w \quad (14.7)$$

and

$$\sum_w L_{ij}^w = L_{ij} \quad \forall i, j \quad (14.8)$$

The above constraints are self-explanatory, and they make sure that the traffic flows on each lightpath are conserved. In addition, they evaluate the number of lightpaths between node pair (i, j) , viz., L_{ij} . The constraints apply if the traffic is either multipoint, or point-to-point.

- *Multicast session topology constraints:*

$$\sum_{i, i \neq s} Z_{is}^{a,d} = 0 \quad \text{and} \quad \sum_{j, j \neq d} Z_{dj}^{a,d} = 0 \quad \forall a, d \in D_a \quad (14.9)$$

$$\sum_{j, j \neq s} Z_{sj}^{a,d} = 1 \quad \text{and} \quad \sum_{i, i \neq d} Z_{id}^{a,d} = 1 \quad \forall a, d \in D_a \quad (14.10)$$

$$\sum_{i, i \neq x} Z_{ix}^{a,d} = \sum_{j, j \neq x} Z_{xj}^{a,d} \quad \forall a, d \in D_a, x, (x \neq s, d) \quad (14.11)$$

The above constraints are similar to the lightpath constraints, except that they apply to the multicast tree traffic, and the continuity is over the lightpaths, and not on the physical links.

Since two or more multicast sessions in the same destination set can share a lightpath to reach their respective destinations, the bandwidth along the shared path could be shared too. The above constraints ensure that the bandwidth of all lightpaths from i to j have not exceeded the physical capacity. The use of the M_{ij}^a variables will avoid multiple counting of the same bandwidth m_a used by several destinations of the same session, and on the same lightpath.

$$M_{ij}^a \geq \sum_{d \in D_a} Z_{ij}^{a,d} / Q \quad \forall a, i, j \quad (14.12)$$

$$M_{ij}^a \leq \sum_{d \in D_a} Z_{ij}^{a,d} \quad \forall a, i, j \quad (14.13)$$

- *Capacity constraints*

$$\sum_{a=1}^K m_a M_{ij}^a \leq L_{ij} * g \quad \forall i, j \quad (14.14)$$

The above ILP produces the locations of the LTEs, and also provisions sessions, in terms of routes and wavelength assignment.

Other Exact Approaches

Other exact approaches were introduced in [5, 9] and [8]. The model in [9] considered multicast traffic grooming on a mesh network, where the objective was to reduce the number of wavelength links. However, the model in [9] assumed that multicast sessions were routed in advance, which removes the routing problem from the formulation. In addition, wavelength conversion

was used, which removes the constraints on wavelength continuity. Moreover, optical splitting can be used. The problem therefore reduced to a bin packing problem on one link at a time.

Using the translucent node architecture shown in Fig. 14.3, the authors in [5] introduced an exact approach, in the form of an Integer Nonlinear Programming formulation, for the network design and session provisioning. The objective function was to also minimize the cost of the electronic equipment. The formulation bears resemblance to the above one, except in one aspect, which is used to determine whether or not an add/drop port is required at a node. This involves terms, for each wavelength, and for each pair of links having a common node, which are the products over all sessions of other terms which are 1 if an add/drop is not needed for each of these sessions. Using a similar architecture, the authors in [8] also introduced an ILP formulation for multicast traffic grooming on ring networks with one traffic unit per session.

The above multicast models assume that all destinations within a destination set have the same traffic requirements. However, different modes of operation of multicasting service include:

- Multicasting with *partial receiver reachability*, in which a subset of the receivers must receive the traffic originating from the source, while every attempt should be made to deliver the traffic to the remaining receivers, if possible, and without increasing the network cost. This corresponds to applications in which users are prioritized, and lower priority users can be accommodated only if this imposes no additional cost. It also corresponds to the case in which the network is designed under tight budget constraints, and users will have to be prioritized in terms of which user is to be served under the given constraints.
- Multicasting in which *traffic may be pruned, or thinned* while propagating downstream, depending on the receivers' needs. Applications of this type include multi-layer coded video streams, in which different receivers may require different stream qualities.

The above two cases were also considered in [7], in which the above ILP was modified in order to accommodate these two cases. The destination set for session a , D_a , is partitioned into two classes, D'_a , and D''_a , such that $D_a = D'_a \cup D''_a$, and $D'_a \cap D''_a = \phi$. All destinations $d' \in D'_a$ have a traffic requirement of m'_a , while destinations $d'' \in D''_a$ have a traffic requirement of m''_a .

For the partial destination reachability case, $m'_a = m''_a$, but D''_a can be reached only if this is not going to increase the network cost. The objective function in this case becomes:

$$\text{Minimize} : \alpha \cdot \sum_n T_n - \beta \cdot \sum_a \sum_{d \in D''_a} \sum_{j, j \neq s} Z_{sj}^{a,d} \quad (14.15)$$

The weights in the above objective function, α and β , should be chosen such that $\alpha \gg \beta$.

Also, for the traffic thinning case, $m'_a \neq m''_a$, and usually $m'_a > m''_a$. The objective function remains the same as in Equation (14.1). However, the constraints must be revised by including a binary variable for each session a on each lightpath, (i, j) , in order to select the correct traffic level, depending on the requirements of the downstream destinations. That is, the traffic delivered on a lightpath used by a session, is by default equal to the m''_a . If the lightpath is used to deliver a traffic level of m'_a , which will cause a binary variable to be equal to 1, a constant quantity equal to $m'_a - m''_a$ is added through multiplying it by the above binary variable.

14.2.4 Static Multicast Traffic Grooming: Heuristic Approaches

Unicast traffic grooming is known to be an NP-hard problem [10, 11]. Multicast traffic grooming is therefore an even harder problem. Therefore, the above exact approaches are only useful for use with small networks, and with limited traffic levels. For large networks, and for higher levels of traffic, other approximate approaches are therefore needed. In addition, the exact approach is still useful for assessing the accuracy of the approximate approaches. Several heuristic approaches were proposed in the literature and will be reviewed here.

Grooming with Lightpath Replacement (GLR)

Reference [7] introduced a heuristic approach based on some observations made from the exact solution of small sized examples. It was found that many sessions are routed along non-shortest paths, and several destinations were reached through lightpaths carrying traffic for such sessions. Those lightpaths were provisioned in a way that makes use of LTEs which were already in place. This observation can be made by inspecting the examples in Figure 14.1.

To illustrate the concept of exploring other trees which can reduce the number of LTEs, consider the example in Fig. 14.4.(a). Two lightpaths were established to deliver traffic (which belong to different sessions) from nodes A to B and C. The grooming factor, g , is 2 units, and the bandwidth requirement per destination is 1 unit. This results in using 4 LTEs. However, when remov-

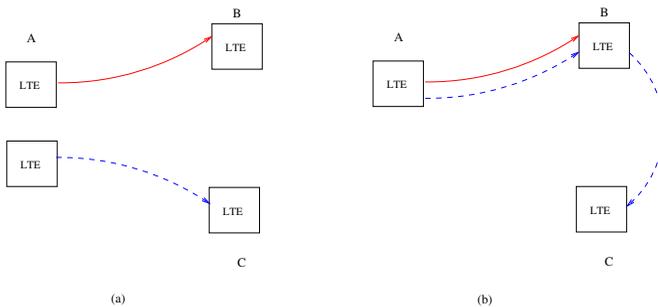


Fig. 14.4. An example to show the two steps of the heuristic

ing the link (A,C), C can be serviced using two branches from A to B, and then from B to C, hence requiring only 3 LTEs, as shown in Figure 14.4.(b).

Since the search space can be very large, the step of exploring other routes and wavelengths is performed by inspecting one lightpath at a time, removing one link on the lightpath, and inspecting whether using existing LTEs can lead to reaching the affected destination while using fewer LTEs.

Heuristics Using Translucent Node Architecture

Reference [5] introduced three heuristics. Although the heuristics have been introduced for networks using a combination of optical and electronic duplication, the algorithms can still work with electronic duplication only.

k-Shortest Path Trees (*k*-SPT)

The first algorithm is the *k*-Shortest Path Tree (*k*-SPT) where the *k*-SPTs are constructed by constructing the SPT, and then removing the links from this SPT, one at a time, in order to obtain the remaining $k - 1$ SPTs.

An example is shown in Fig. 14.5, where there are two sessions: the first one is $\langle A, \{C, G\}, 1 \rangle$, while the second one is $\langle A, \{E, G\}, 1 \rangle$. The grooming factor, g , is 2. Figure 14.5.(a) shows the routing under SPT for each session, which requires 10 LTE ports, which are shown as squares. However, in Figure 14.5.(b), optical splitting is used at node B, which reduces the number of LTEs to 5. Note that in Fig. 14.5.(b), traffic from the second session is delivered to node C, which is not part of the destination set of this session. This is due to the optical splitting implemented at node B.

Grooming with Rerouting Sessions (GRS)

The second algorithm in [5] is called Grooming with Rerouting Sessions (GRS). In this algorithm, if a session cannot be routed on its shortest path tree on already used wavelengths due to the unavailability of resources on some links, such links are treated as bottleneck links. If the number of such bottleneck links does not exceed two links, then the session is rerouted to

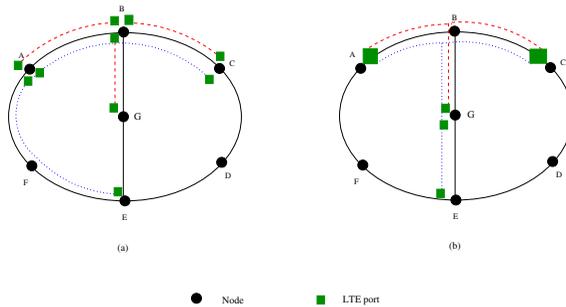


Fig. 14.5. An example of the *k*-SPT heuristic

avoid those links. Otherwise, the next available wavelength is assigned to the shortest path tree.

Grooming by Computing Overlapped Trees (GCOT)

The third and last algorithm is Grooming by Computing Overlapped Trees (GCOT). This algorithm tries to pack sessions onto wavelengths in a greedy manner.

Performance Evaluation

It was shown in [5] that the k-SPT algorithm outperforms both the GRS and GCOT algorithms, both in terms of the number of LTEs and the number of wavelength channels. Also, as the value of k increases, the performance of k-SPT increases further, but of course at the expense of an added complexity. By comparing the performance of the algorithms to the optimal solution obtained by solving the nonlinear program mentioned earlier, it was shown that the results of the algorithm are within 66% of the optimal. However, in most cases they are within 10–20% of the optimal.

The GLR algorithm described above is very much related to the k-SPT algorithm, since it also investigates several shortest paths. However, there are two differences between the two algorithms. First, under GLR, the number of alternate shortest paths is not fixed, while it is fixed under k-SPT. Second, under k-SPT, shortest path trees are constructed after removing a link, while under GLR, the original tree is augmented in order to reach the destinations which have been disconnected.

Multicast Traffic Grooming with Sparse Optical Splitters

Reference [9] considered a network with Multicast Capable (MC) nodes (equipped with optical splitters) and Multicast Incapable (MI) nodes. It is assumed that each MI node has at least one adjacent MC node as well as each node is equipped with a bank of wavelength converters. The paper presented a two phase algorithm for provisioning multicast traffic sessions, namely, *Multicast Tree Construction (MTC)* and *First Fit Grooming Algorithm (FFGA)*. MTC constructs shortest path trees from each MC node, and allows an MI node to use an adjacent MC node as a virtual source, or root. The MI node, therefore, routes its multicast traffic to that adjacent MC node, and the MC node will split the traffic and route it to the destinations. The FFGA is a first fit greedy algorithm that allocates traffic streams on a link to the first available wavelength channel on that link. If no wavelength channel is available, a new wavelength channel is created.

14.2.5 Dynamic Multicast Traffic Grooming

With the use of IP over WDM, MPLS over WDM, or with NGS, traffic sessions now tend to exhibit a dynamic nature, as opposed to the static

nature assumed in the previous section. Sessions arrive according to a certain arrival process, and they are characterized by holding times, which are taken from a certain distribution. Since it is practically impossible to design optical networks such that they accommodate all such dynamic sessions, most efforts have concentrated on devising call acceptance, and session provisioning strategies that will try to reduce the session blocking probabilities.

A number of session provisioning strategies for multicast traffic were introduced in the literature, and will be presented below.

The Maximizing Minimum Freeload (MMFL) Algorithm

Reference [12] introduced a session provisioning strategy for dynamic multicast traffic, assuming translucent nodes with the architecture shown in Fig. 14.3. The objective of this algorithm is to increase resource utilization, and to minimize the blocking probability for future arriving requests. This is done by using paths which will maximize the bandwidth capacity left after routing a multicast tree. In other words, the problem can be considered as a max-min problem, where the minimum left over bandwidth on all links is maximized.

The bandwidth capacity left after routing the tree is called the *freeload*, and the freeload on link (i, j) which has an available capacity $c_{i,j}$ on wavelength w after routing a session with bandwidth requirement B on this link, and this wavelength, is given by

$$\frac{c_{i,j} - B}{g} \quad (14.16)$$

where g is the grooming factor. The minimum freeload on all wavelengths is calculated assuming that the session is provisioned, and the routing that yields the maximum over these minima is used.

The algorithm introduced in [12] is called *Maximizing Minimum Freeload* (MMFL). It is assumed that a freeload graph for each wavelength, w , is formed for each session. The current such graph is referred to as G_w , and the new graph, G'_w , is formed for each session by applying Equation (14.16) on all links of the graph using the session requirement. After deciding on the routing and wavelength assignment of the session, only G_w on which the session is routed is updated by discounting the allocated link capacities. MMFL is executed for every arriving multicast session.

Figure 14.6 shows an example of the MMFL in which there are two sessions, the first is $\langle A, \{C, G\}, 1 \rangle$, while the second is $\langle A, \{C, E\}, 1 \rangle$. Both sessions have bandwidth requirements of one unit, while $g = 2$. It is assumed that there is a single wavelength in the network. In Fig. 14.6.(a), session 1 arrives, and the freeload factor for all fibers is calculated as $\frac{2-1}{2} = 0.5$. Session 1 is

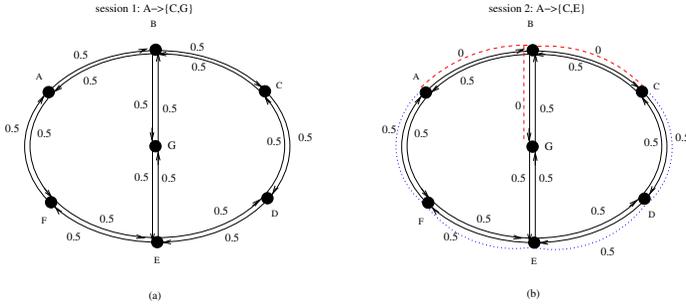


Fig. 14.6. An example of the operation of the MMFL algorithm

routed on the SPT shown by the dashed lines in Figure 14.6.(b). When session 2 arrives, in Fig. 14.6.(b) the freeload factors for session 2 are calculated as shown in the figure. Fibers used by session 1 have a freeload factor of $1 - 1/2 = 0$, while all other fibers have a freeload factor of 1. Session 2 is provisioned on the SPT shown by the dotted lines in Fig. 14.6.(b). If there was another wavelength in the network, the freeload diagram for the second wavelength, when session 2 arrives, would be identical to that in Fig. 14.6.(a), and session 2 would be provisioned similar to that in Fig. 14.5.(b), but on the second wavelength.

It was shown in [12] that MMFL outperforms both fixed and adaptive SPT provisioning techniques in terms of call acceptance ratio, since MMFL attempts to uniformly distribute the session loads over wavelengths, hence resulting in a better chance of call acceptance. The same observation also holds for resource utilization.

It is to be noted that this algorithm does not consider the effect of using optical splitting on accommodating sessions. For example, the route provisioning shown in Fig. 14.5.(b) will not be selected by this algorithm.

Sequential and Hybrid Multicast Traffic Grooming

References [13, 14] assumed a node architecture similar to that in Fig. 14.3, except that duplication is always done in the optical domain and that multicasting is always supported on a light-tree. All traffic on the same light-tree will be delivered to the same set of destinations on the light-tree, whether they are part of the destination set or not.

Multicast Grooming Approaches

In order to support dynamic traffic, once a multicast session is to be accommodated, one of four traffic grooming approaches may be used. These approaches are described below, with the aid of the example given in Fig. 14.7, assuming that there is one wavelength channel and all sessions have a traffic requirement of 1 unit, while $g = 2$. It is also assumed that sessions 1, 2 and 3 are already established on the network.

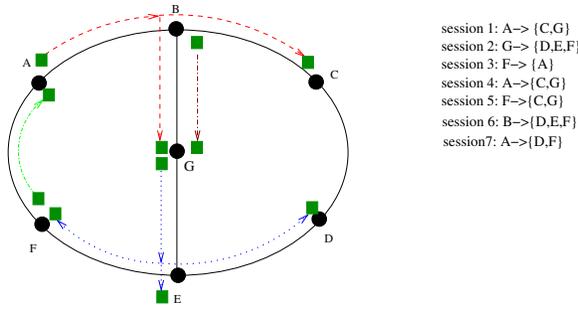


Fig. 14.7. Examples of the approaches used in sequential multicast traffic grooming

Single-Hop Provisioning: In this approach, an existing light-tree with available bandwidth to support the new session is used to groom the multicast session at the logical layer. The new multicast session will be served using a single-hop logical route. Note that for this approach, the new multicast session must have the same source and multicast destinations as those associated with the logical route. For example, in Fig. 14.7, when session 4 arrives, it can be accommodated on the light-tree used by session 1.

Multi-Hop Provisioning: A multicast session can be provisioned on the logical topology by routing data on more than one light-tree (only two hops are allowed here).

In this approach, an existing light-tree whose destinations are the same as those of the new multicast session (called the *to-destinations light-tree* (TDLT)), but with traffic from a different source, is used with combination of a single-hop lightpath whose source is the same as the source of the new session and its destination is the source of the TDLT (called *from-source light-path* (FSLP)). The new multicast session is served on the combination of FSLP and TDLT. The TDLT and FSLP must have enough capacity to accommodate the arriving session. For example, when session 5 arrives, it can be provisioned using an FSLP F to A (which is used to accommodate session 3), and then TDLT from A to C and G (which is used to accommodate session 1).

Hybrid Provisioning: A session can be routed/provisioned over a combination of existing lightpaths or light-trees (logical routing) and a newly created light-path or light-tree (physical routing, e.g., RWA/MC-RWA). In this approach, the optical layer must keep updated databases about the connectivity of both the logical and physical topologies as well as the resource utilization across both layers. The idea of a hybrid provisioning approach is defined to find a combined route of an existing logical segment (light-tree) and an un-provisioned physical segment (lightpath) that needs to be set up.

For example, when session 6 arrives, the light-tree used by session 2 from G to {D,E,F} can be used to reach F, E and D from G. In addition, a light-path from B to G is needed, and is therefore provisioned.

Non-Restricted Multi-Hop Provisioning: This approach allows light-trees to carry traffic destined to only a subset of the destinations of the light-tree (this includes an empty subset, which corresponds to the case in which the light-tree is used as a first hop). For example, when session 7 arrives, the signal can be transmitted on the light-tree used by session 1. Since none of the destinations of session 7 are on this tree, at node G the signal is carried on the light-tree used by session 2 to reach D and F. Notice here that the signals also reach nodes C, and E, which are not in the destination set of session 7.

Traffic Grooming Heuristics

Using the above four approaches, the authors introduced four heuristics for provisioning multicast calls with subwavelength traffic rates.

Logical-First Sequential Routing (LFSR): The network first tries to accommodate the call on the logical topology making use of the already existing connections. Depending on the grooming approach to be used (i.e., Single-Hop or Multi-Hop), if no available single/multi-hop route were found on the logical topology, then a new light-tree destined to all the multicast destinations is set up on the physical topology.

Physical-First Sequential Routing (PFSR): The network attempts to accommodate a session on the physical layer first. If the new light-tree is established successfully, a new logical/virtual route (light-tree) is created in the logical layer. If the physical routing fails, then single or multiphop routing on the logical topology is attempted.

Logical-First Hybrid Routing (LFHR): The hybrid provisioning approach is combined with sequential provisioning approaches to achieve a grooming scheme that is biased toward serving multicast traffic demands in comparison with all other sequential grooming approaches. Specifically, the *LFSR* is invoked to find an existing lightpath to provision an FSLP to deliver the traffic from the source of the session to the root of a TDLT that reaches the destinations. If this fails, a new light-path is created and used to reach the source of the TDLT.

Non-Restricted Logical-first Sequential Routing (NRLFSR): Similar to *LFSR*, but a non-restricted multi-hop provisioning approach is used when *LFSR* fails.

Performance Evaluation

Using a simulation study, and assuming that only one third of the calls are multicast, it was shown in [14] that among the restricted algorithms (*LFSR*, *PFSR* and *LFHR*), *PFSR* has the smallest blocking probability. The use of multihop session provisioning reduces the blocking probability even further,

and provides a reduction in the blocking probability that can be very close to 20%. On the other hand, it was also shown that the non-restricted routing always outperforms constrained routing at light load. The last observation is expected since the routing problem has fewer constraints. However, as the load is increased, the restricted LFSR outperforms NRLFSR, since it is more conservative in using bandwidth to accommodate new sessions. When most of the sessions are multicast sessions, the non-restricted routing approach prevails.

14.2.6 Multicast Traffic Grooming: Summary

In Table 14.2, we provide a summary of the above multicast traffic grooming protocols. For each protocol, the duplication domain for which the algorithm was designed, the mode of operation, and the complexity of the algorithm are indicated.

14.3 Many-to-One Traffic Grooming

In this section, we illustrate the concept of many-to-one traffic grooming, and how aggregating traffic, if possible, can result in resource reduction, with the help of a simple example. Figure 14.8 shows a network consisting of 7 nodes. Suppose that there is only one session which consists of 4 sources, S_1 , S_2 , S_3 and S_4 , located at nodes 1, 3, 4 and 6, respectively. The destination of this session, d , is located at node 7. Let the capacity of a single wavelength be 48 units and the traffic originating from each source be 24 units. We refer to the traffic originating from each source as a stream. Without traffic grooming, the session $\{S_1, S_2, S_3, S_4 \rightarrow d\}$ must be accommodated using 3 wavelengths and 8 LTEs (2 LTE for each source-destination pair). However, with traffic grooming, the number of LTEs and wavelengths can be reduced

Table 14.2. Summary of multicast traffic grooming heuristic algorithms

Algorithm	Duplication domain	Operation	Complexity
GLR [7]	electronic	Static	$O(K \cdot M \cdot N^2 \log_2 N)$
k-SPT [5]	electronic-optical	Static	$O(K \cdot k \cdot N^2 \log_2 N)$
GRS [5]	electronic-optical	Static	$O(K^2 \cdot M \cdot N^2 \log_2 N)$
GCOT [5]	electronic-optical	Static	$O(K^2 \cdot M \cdot W)$
MTC+FFGA [9]	optical	Static	$O(K \cdot N^2 \log_2 N + K \cdot M \cdot W)$
MMFL [12]	electronic-optical	Dynamic	$O(K \cdot N^2 \log_2 N)$
LFSR [14]	optical	Dynamic	$O(W \cdot N^2)$
PFSR [14]	optical	Dynamic	$O(W \cdot N^2)$
LFHR [14]	optical	Dynamic	$O(W \cdot N^2)$
NRLFSR [14]	optical	Dynamic	$O(W \cdot N^2)$

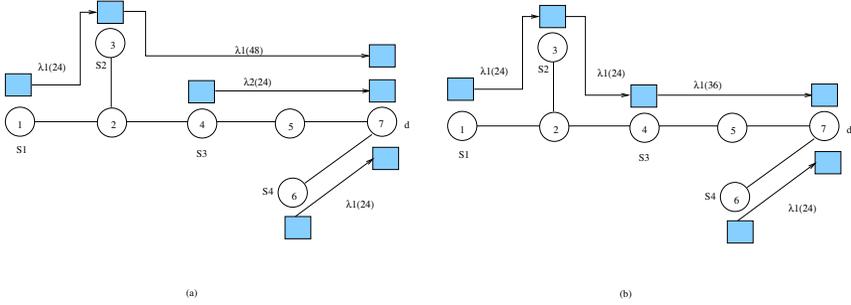


Fig. 14.8. An example of many-to-one traffic grooming: (a) without aggregation and traffic reduction; (b) with aggregation and traffic reduction

to 7 and 2, respectively, as shown in Fig. 14.8(a). If the streams from different sources destined to the same destination, d , may be aggregated on their way, the resources can be further reduced. If we assume that every time traffic is aggregated, 50% of the aggregated traffic will be dropped, then streams from S_1 and S_2 are first aggregated at node 3 and then streams from S_1 , S_2 and S_3 are aggregated at node 4. This accommodation requires only 1 wavelength and 6 LTEs, as shown in Fig. 14.8.(b). Therefore, many-to-one traffic grooming with aggregation can result in a substantial reduction in the cost of the network in terms of the number of LTEs and the number of wavelengths.

14.3.1 Network Model

Similar to the multicast traffic grooming case, we consider a network with N nodes, and an arbitrary mesh topology. Each link corresponds to two bidirectional fibers in opposite directions, where each fiber carries W wavelengths, and the grooming ratio is g . We allow each many-to-one session to traverse multiple lightpath hops, while each lightpath hop itself may span multiple physical links.

Each session, a , will have one destination, d_a , and a set of sources, S_a , with each source $s \in S_a$ generating a number of basic traffic units, $m_{a,s}$. The stream from s to d in session a is denoted by $\phi_{s,d,a}$. Traffic from different sources in the same session may be aggregated such that if f streams from session a are aggregated, each stream contributes a fraction r_f^a of its traffic.

14.3.2 Static Many-to-One Traffic Grooming with Aggregation

We address the many-to-one traffic grooming problem such that, given a traffic matrix, all the traffic should be accommodated with the least number of LTEs. We also consider the situation in which traffic can be aggregated using arbitrary, but application dependent, aggregation ratios. Aggregation

can be used to reduce the amount of traffic that has to be carried by the network, and hence possibly reduce the amount of resources. We assume that aggregation, when performed, will be implemented at layer 1. We expect that the most flexible way of implementing many-to-one traffic grooming would be by using DoS switches, and NGS protocols, e.g., by delineating and flagging data blocks which may be dropped during aggregation. GFP extension headers [15] can facilitate this implementation. This area of research is very new, and all the approaches presented in this section are based on reference [16].

Exact Approach: Mixed Integer Linear Program (MILP) Formulation

The optimal formulation for the network and connection provisioning problem, such that the number of LTEs is minimized, is based on a Mixed Integer Linear Program (MILP). Several of the variables and constraints defined for the multicast traffic grooming problem of Section 14.2.3 are valid for the many-to-one traffic grooming problem. In addition to the parameters defined in the network model above, and those defined in Table 14.3, the MILP uses the following parameters:

Objective function:

$$\text{Minimize : } \sum_n T_n \quad (14.17)$$

Subject to:

- *Number of LTEs:*
The constraints are similar to those in Equations (14.2) and (14.3).
- *Lightpath level constraints:*
The constraints are also similar to those in Equations (14.4), (14.5), (14.6), (14.7) and (14.8).

Table 14.3. Many-to-one traffic grooming MILP basic parameters

$Z_{ij}^{a,s}$	a real number between 0 and 1, which takes non zero values if and only if the traffic stream from source $s \in S_a$, is using a lightpath from i to j
M_{ij}^a	a binary indicator; is 1 if and only if at least one of the sources of session a is using a lightpath from node i to j to reach the destination, i.e., $\exists s \in S_a$, such that $Z_{ij}^{a,s} = 1$
$I_{ij}^{c,a,f}$	a binary indicator; is 1 if and only if the number of sources $s \in S_{c_a}$, on a lightpath from node i to j , is $\geq f$
X_{ij}^a	a real number which represents the amount of traffic on a lightpath from node i to j due to all sources in S_a
$J_{ij}^{a,b}$	a binary indicator; is 1 if and only if session a and b are groomed on the same lightpath from i to j
$Y_{ij}^{a,b}$	a real number and is a product of $J_{ij}^{a,b}$ and X_{ij}^b

- *Session topology constraints:*

$$\sum_i Z_{is}^{a,s} = \sum_j Z_{dj}^{a,s} = 0 \quad \forall a, s \in S_a \quad (14.18)$$

$$\sum_{j,j \neq s} Z_{sj}^{a,s} = \sum_{i,i \neq d} Z_{id}^{a,s} = 1 \quad \forall a, s \in S_a \quad (14.19)$$

$$\sum_{i,i \neq x, i \neq d} Z_{ix}^{a,s} = \sum_{j,j \neq x, j \neq s} Z_{xj}^{a,s} \quad \forall a, s \in S_a, x, (x \neq s, d) \quad (14.20)$$

The above constraints conserve traffic flows on the lightpaths from the sources to the destination, in each of the sessions.

Moreover, the following constraints will ensure that once data streams, from more than one source, are aggregated at some node, they will not split again until they reach the destination.

$$M_{ij}^a \geq \sum_{s \in S_a} Z_{ij}^{a,s} / Q \quad \forall a, i, j \quad (14.21)$$

$$M_{ij}^a \leq \sum_{s \in S_a} Z_{ij}^{a,s} \quad \forall a, i, j \quad (14.22)$$

$$\sum_j M_{ij}^a \leq 1 \quad \forall a, i \quad (14.23)$$

To model aggregation, and to determine the amount of traffic before and after aggregation on each of the lightpaths, the following constraints are used to set the binary indicator $I_{ij}^{a,f}$ on the lightpaths from i to j to one if and only if the number of sources $s \in S_a$ is greater than or equal to f .

$$\sum_{s \in S_a} Z_{ij}^{a,s} = \sum_{f=1}^{|S_a|} I_{ij}^{a,f} \quad \forall a, i, j \quad (14.24)$$

$$I_{ij}^{a,f} \leq I_{ij}^{a,f-1} \quad 2 \leq f \leq |S_a|, \forall a, i, j \quad (14.25)$$

Once the $I_{ij}^{a,f}$ variables have been determined by the above constraints, they are used to determine the exact amount of traffic after aggregation on lightpath(s) from node i to node j due to the sources in S_a . This is done using the following set of constraints.

$$X_{ij}^a \geq r_f^a * \sum_{s \in S_a} m_{a,s} Z_{ij}^{a,s} - Q \sum_{s \in S_a} Z_{ij}^{a,s} + Q * f * I_{ij}^{a,f} \quad 1 \leq f \leq |S_a|, \forall a, i, j \quad (14.26)$$

$$X_{ij}^a \leq r_f^a * \sum_{s \in S_a} m_{a,s} Z_{ij}^{a,s} + Q \sum_{s \in S_a} Z_{ij}^{a,s} - Q * f * I_{ij}^{a,f} \quad 1 \leq f \leq |S_a|, \forall a, i, j \quad (14.27)$$

The above constraints will compute the exact amount of traffic after aggregation on lightpath(s) from node i to node j due to the sources in S_a , since when $f < \sum_{s \in S_a} Z_{i,j}^{a,s}$ or $f > \sum_{s \in S_a} Z_{i,j}^{a,s}$, Equations (14.26) and (14.27) will evaluate X_{ij}^a to be in the range $[-Q, +Q]$. However, when $f = \sum_{s \in S_a} Z_{i,j}^{a,s}$, X_{ij}^a will evaluate to the correct amount of traffic after aggregation.

- *Capacity constraint:*

The following constraint bounds the total capacity on lightpath(s):

$$\sum_{a=1}^K X_{ij}^a \leq L_{ij} * g \quad \forall i, j \quad (14.28)$$

Similar to the multicast traffic grooming problem, the above MILP will determine the locations of the LTEs, as well as the provisioning of sessions.

Heuristic Approach

Reference [16] has also presented a heuristic approach for many-to-one traffic grooming on random topologies, which also performs data aggregation. The heuristic is based on a Dynamic Programming (DP) style algorithm, which is referred to as the Many-to-one Traffic Grooming heuristic based on Dynamic Programming (MTG-DP). That is, the algorithm proceeds in stages, where in stage i different subsets of i streams each are routed and provisioned. The number of subsets, as will be explained below, is equal to the total number of streams. In stage i , stream j is provisioned together with $i - 1$ streams, where such streams are chosen as one of the subsets of streams provisioned in stage $i - 1$, such that the provisioning cost is minimized. This makes the computations at stage i dependent on stage $i - 1$ only, hence the dynamic programming style.

An example of the DP-style algorithm is shown in Fig. 14.9. If each many-to-one session, a , consists of $|S_a|$ streams, one from each source $s \in S_a$, and if Φ is the total number of streams from all the many-to-one sessions, i.e., $\Phi = \sum_a |S_a|$, then, as shown in the figure, the algorithm consists of Φ stages, while each stage further consists of Φ steps. At each step of a stage, a new stream is accommodated (i.e., the stream is assigned a specific route and a specific wavelength) after conducting a systematic search, such that the decisions at each step of a stage depend only on the previous stage. For the systematic search at each stage, all the assignments of the previous stage are examined and the assignment that can accommodate the current stream with the least cost is determined. In Fig. 14.9, steps 2 and 4 of stages 2 and 3 are shown in detail. The outcome of each step is a selection of a set of accommodated streams, which are called an assignment. At the end of each stage, a collection of several best assignments are available, and after the final stage, the assignment with the least cost is selected.

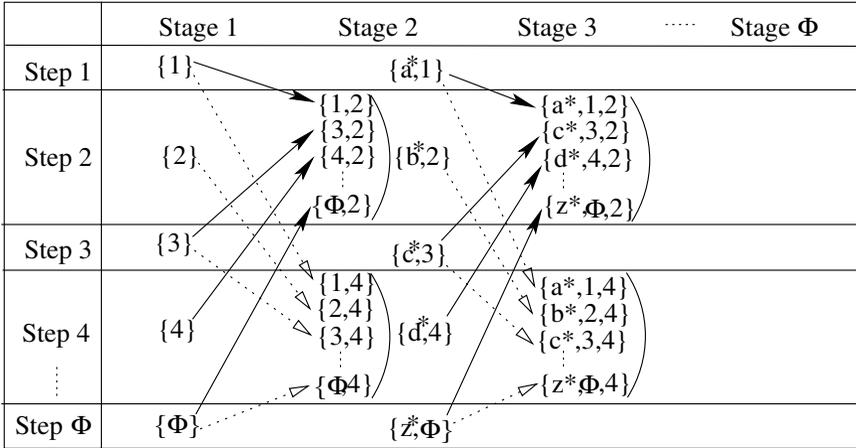


Fig. 14.9. Graphical depiction of the many-to-one traffic grooming heuristic. It is assumed that $a^* \neq 1, b^* \neq 2, c^* \neq 3, d^* \neq 4,$ and $z^* \neq \Phi$

The complexity of this heuristic was shown in [16] to be $O(\Phi^3 N^2 \sum_{c_k} |S_{c_k}|^2)$. It was also shown, through comparison to the exact approach, that this heuristic produces results which are between 10% and 38% of the exact solution.

14.4 Research Issues in Multipoint Traffic Grooming

Multipoint traffic grooming is a new field of research, and there are several open research problems in this field. As was evident from the discussion, the complexity of the optimal network design and session provisioning problem exceeds that of the normal unicast traffic grooming, which is known to be NP-hard. Therefore, approximate and heuristic approaches need to be developed for use with large networks, and increased traffic levels. In addition, session provisioning of multipoint traffic sessions under dynamic traffic conditions is an important problems, especially with the use of IP and MPLS over WDM. Tractable and efficient procedures are therefore required. The development of accurate performance models for the evaluation of blocking probabilities under this type of traffic is still an open area of research.

Finally, the survivable multipoint traffic grooming problem, where multipoint low-speed connections need to be protected, is still not receiving much attention in the literature. Recently, some work addressed the traffic grooming problem for survivable WDM networks; however only unicast traffic demands were considered. There is a need to propose and develop new approaches to provision and protect low-speed multipoint traffic demands.

References

1. B. Quinn and K. Almeroth, "IP multicast applications: Challenges and solutions." Internet Engineering Task Force (IETF) Request for Comments (RFC) 3170, Sept. 2001.
2. G. Rouskas, "Optical layer multicast: Rationale, building blocks, and challenges," *IEEE Network*, vol. 17, no. 1, pp. 60–65, 2003.
3. A. Hamad, T. Wu, A. E. Kamal, and A. K. Somani, "On multicasting in wavelength-routing networks," *Computer Networks*, vol. 14, pp. 3105–3164, Nov. 2006.
4. A. Hamad and A. E. Kamal, "A survey of multicasting protocols for broadcast-and-select single-hop networks," *IEEE Network*, vol. 16, pp. 36–48, July 2002.
5. G. Chowdhary and C. S. R. Murthy, "Grooming of multicast sessions in WDM mesh networks," in *Workshop on Traffic Grooming*, 2004.
6. D. Cavendish et al., "New transport services for next-generation sonet/sdh systems," *IEEE Communications*, vol. 40, pp. 80–87, May 2002.
7. R. Ul-Mustafa and A. E. Kamal, "Design and provisioning of WDM networks with multicast traffic grooming," *IEEE Journal on Selected Areas in Communications, Part II: Optical Communications and Networking*, vol. 24, pp. 37–53, Apr. 2006.
8. H. V. Madhyastha, G. V. Chowdhary, N. Srinivas, and C. Siva Ram Murthy, "Grooming of Multicast Sessions in Metropolitan WDM Ring Networks," *Computer Networks*, vol. 49, no. 4, pp. 561–579, 2005.
9. A. R. Billah, B. Wang, and A. A. Awwal, "Multicast traffic grooming in WDM optical mesh networks," in *Proceedings of the Conference on Global Communications (GLOBECOM)*, 2003.
10. A. L. Chiu and E. H. Modiano, "Traffic grooming algorithms for reducing electronic multiplexing costs in WDM ring networks," *IEEE Journal of Lightwave Technology*, vol. 18, pp. 2–12, Jan. 2000.
11. R. Dutta, S. Huang, and G. Rouskas, "On optimal traffic grooming in elemental network topologies," in *the Proceedings of Opticomm*, pp. 13–24, 2003.
12. G. Chowdhary and C. S. R. Murthy, "Dynamic multicast traffic engineering in WDM groomed mesh networks," in *Workshop on Traffic Grooming*, 2004.
13. A. Khalil, C. Assi, A. Hadjiantonis, G. Ellinas, and M. Ali, "On multicast traffic grooming in WDM networks," in *IEEE Int. Symp. on Computers and Communications*, pp. 282–287, 2004.
14. A. Khalil, A. Hadjiantonis, G. Ellinas, and M. Ali, "Sequential and hybrid grooming approaches for multicast traffic in WDM networks," in *Proceedings of the Conference on Global Communications (GLOBECOM)*, 2004.
15. E. Hernandez-Valencia, M. Scholten, and Z. Zhu, "The Generic Framing Procedure (GFP): An Overview," *IEEE Communications*, vol. 40, no. 5, pp. 63–71, May 2002.
16. R. Ul-Mustafa and A. E. Kamal, "Many-to-one Traffic Grooming in WDM Networks with Variable Aggregation Ratios," *IEEE Journal on Selected Areas in Communications, Part II: Optical Communications and Networking*, vol. 24, pp. 68–81, Aug. 2006.

Part III

Frontiers

Multi-Domain Traffic Grooming

Nasir Ghani, Qing Liu, Nageswara S. V. Rao, Ashwin Gumaste

15.1 Introduction

The last decade has seen many advances in high-speed networks. At the fiber level, dense wavelength division multiplexing (DWDM, Layer 1) has gained favor as a terabit solution, capable of “lightpath” circuit routing. Meanwhile next-generation SONET/SDH (NGS) (Layer 1.5) [1] has gained strong traction in the metro/edge, providing flexible “sub-rate” aggregation. Finally, Ethernet and IP networks (Layers 2, 3) have seen new quality of service (QoS) provisions via the differentiated services (Diff-Serv) and integrated services (Intserv) frameworks. Related control standards have also emerged, most notably multi-protocol label switching (MPLS) for Layer 2/3 flow-based QoS support. Further generalizations have even adapted this solution for “non-packet-switching” layers, i.e., generalized MPLS (GMPLS) [2].

As the above solutions undergo widespread deployment, a complex interconnection of different data-planes has resulted. From a theoretical aspect, this is embodied by an interconnection of multiple horizontal domains (i.e., networks) comprising of different vertical layers (i.e., technologies), as shown in Fig. 15.1 [3]. These segmentations are based upon various factors, such as technology, scalability, geographic, economic, administrative, etc. Moreover as the scale and reach of services expands, there is a pressing for “end-to-end” provisioning across heterogeneous domains, i.e., horizontal-vertical control. A good example is the growth in “e-science” applications [3]. However, inter-layer provisioning today is still done via manual provisioning of domain-specific control planes, giving high inefficiency and long lead times, i.e., hours to days.

Multi-domain networking has long been supported in IP and ATM networks. Moreover, new efforts are also introducing these capabilities in optical transport networks. Nevertheless, the broader topic of provisioning across heterogeneous network layers has not been addressed in detail. Note that there is a clear distributed multi-domain grooming aspect to this problem as different circuit/flow granularities are involved, i.e., gigabit wavelengths, “sub-rate”

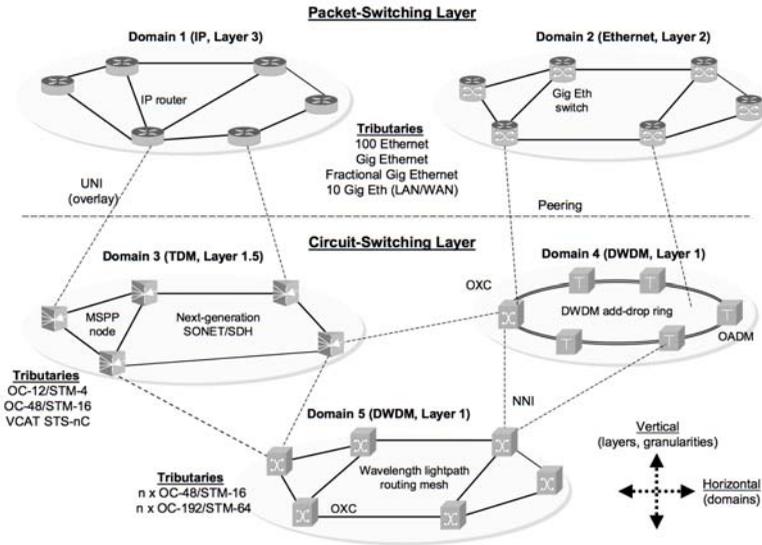


Fig. 15.1. Multi-layer, multi-domain networking infrastructures

SONET/SDH tributaries, MPLS label switched paths (LSP), etc. Overall, this is a very difficult problem owing to some key complicating factors. Foremost, there are no unified control plane standards for integrating multiple layers. In addition related provisioning algorithms are lacking, particularly those operating in distributed (intra, inter-carrier) settings with no/partial global state.

This chapter addresses the emerging area of multi-domain grooming and is organized as follows. First, Section 15.2 presents a brief survey of standards for multi-domain networks. Subsequently Section 15.3 surveys some work on multi-domain protocols and algorithms. Section 15.4 then delves into the relatively unexplored area of multi-domain grooming and details key directions in inter-layer routing, path computation, signaling, and survivability. Finally Section 5 presents overall conclusions and future directions.

15.2 Multi-Domain Standards

Over the years, a wide range of multi-domain networking standards have been developed by various standards organizations, including the IETF, ITU-T, and OIF. These efforts are briefly reviewed here.

15.2.1 ITU-T Standards

The ITU-T has matured a comprehensive automatically switched transport networks (ASTN) framework for multi-domain optical networks. The reference architecture for ASTN is G.8080 [2] and describes a group of components to control transport resources to setup, maintain, and release a client layer connections. However, the internal topology and connectivity of the underlying optical layers is not made visible to the client layer and is instead treated as a sub-network point pool (SNPP) link (“virtual” link). Now ASTN follows a hierarchical design in which designated entities perform control for single and multiple layers. For example, its routing hierarchy consists of areas with requirements for inter-area auto-discovery, auto-provisioning and auto-restoration. However, this model is flexible and each domain’s internal control plane can be tailored to the particular types and capabilities of the equipment within the domain. Nevertheless, this overall framework only addresses architectural design and no explicit protocols specifications are provided. It is here that developments within the IETF and OIF are of crucial importance.

15.2.2 OIF Standards

The OIF has developed several control standards for optical network interfacing, including a client–carrier user-network interface (UNI) and a carrier–carrier network node interface (NNI). The UNI implements bandwidth signaling for client devices (i.e., IP/MPLS routers) to request/release “optical” connections from underlying carrier SONET/SDH or DWDM domains, i.e., “optical dialtone”. Since there is no trust relationship here, carrier topology information is not propagated to the client side, i.e., overlay model [2]. The latest UNI 2.0 features much improved capabilities for security, bandwidth modification, etc. Meanwhile, the NNI interface implements inter-domain functions for reachability/resource exchange and setup signaling. Now two variants have been designed here, interior NNI (I-NNI) and external-NNI (E-NNI). The former interfaces nodes within the same administrative area whereas the latter serves adjacent areas. E-NNI relegates all interfacing issues to domain boundaries, thereby removing restrictions on domain-internal control and equipment interoperability. Furthermore, NNI adopts a link state routing approach, e.g., via single/stacked instances of GMPLS routing protocols.

The overall UNI–NNI combination facilitates rapid deployment of new services across multi-vendor “optical” domains. Moreover, this framework works for both optical DWDM and SONET/SDH domains, as facilitated by the different granularities supported in UNI/NNI signaling. Hence the NNI standard can support multi-layer interfacing at the circuit-switching level. However, since there is no resource exchange across the UNI, higher layer IP/MPLS networks must treat underlying “optical” links as tunneled links. This setup will be problematic for client layer routing since the number of Layer 3 adjacencies will grow per the square of border nodes, i.e., connection

mesh problem. Clearly, as the number of border nodes grows, this routing setup becomes less scalable, considering the amount of adjacency information has to be maintained. Hence clients may have to run data-plane inter-domain routing protocols across UNI-provisioned connection links.

15.2.3 IETF Standards

Well-established IP network architectures comprise a hierarchy of autonomous systems (AS) and routing areas (domains). Within areas, routers run associated interior gateway protocols (IGP) such as open shortest path first (OSPF) or intermediate-system to intermediate-system (IS-IS) [2]. These protocols use link-state routing to maintain link state databases (LSDB). Meanwhile at the inter-AS level, commensurate exterior gateway protocols (EGP) are used. The mainstay here is border gateway protocol (BGP) which runs between exterior gateway nodes and uses distance vector routing to provide end-point reachability exchange. In addition, OSPF also provides an additional level for disseminating “higher-layer” aggregated state between area border routers (ABR). Note that many operators will want to limit internal state dissemination in inter-carrier settings owing to privacy and security concerns.

With growing flow-level QoS requirements in IP networks, new extensions have also been proposed for OSPF. For example, OSPF-traffic engineering (OSPF-TE, RFC 2676) provides extensions and opaque link state (LSA) definitions for “QoS-related” link state to support advanced constraint-based routing (CBR). Meanwhile there have also been proposals for augmenting BGP messaging with added QoS attributes. For example, [4] exchanges QoS-enabled reachability information via new BGP message attributes. However, these offerings do not alter the inherent distance vector design. Finally, the hierarchical routing has also been implemented in older asynchronous transfer mode (ATM) networks via the private network-to-network interface (PNNI) protocol, i.e., via peer group leaders.

Meanwhile the GMPLS framework devises new routing/signaling versions to manage “circuit-switched” entities, e.g., TDM circuits and DWDM light-paths [2]. For example, new OSPF-TE opaque LSA definitions have been added for DWDM and SONET/SDH links. Hence TE databases (TEDB) can now store information fields for wavelengths/usages, timeslots/usages, shared risk link groups (SRLG) diversity, etc. New enhancements have also been added to the reservation protocol, RSVP-TE, to implement DWDM and SONET/SDH circuit setup/takedown, e.g., hard state, recovery, etc. Additionally, RSVP-TE also provides a loose route (LR) feature to specify a partial (high-level) node sequences along with signaling-based explicit route (ER) expansion to resolve the full path route. Finally a new link management protocol (LMP) has been developed for fault discovery and localization.

However, GMPLS does not provide any specific routing extensions for multi-domain BGP as its inherent distance-vector design is not well-suited

for circuit routing. Therefore, two-level OSPF-TE is the most germane inter-domain protocol for GMPLS. Nevertheless, OSPF-TE assumes a peer-to-peer model in which all nodes perform identical routing/signaling. Clearly, today's emerging diversified, multi-layer domains settings will prove problematic here. For example, many vendor solutions may not provide GMPLS support and instead use centralized vendor-proprietary network management systems (NMS) or operations support systems (OSS) and TL-1 messaging [1]. In these cases, proxy solutions can be used to translate external vendor-specific control formats [5].

The IETF is also addressing TE path computation via its path computation element (PCE) architecture [6, 7]. This framework allows path computation clients (PCC) to interact with PCE entities to resolve constraint-based routes. A PCE either resides in a standalone manner or is co-located with a network node and has access to domain-level resource/policy databases. This architecture provides a request/response protocol for PCC-PCE and PCE-PCE interaction, i.e., PCE protocol (PCEP) [7]. Finally, automated discovery allows PCC entities to locate a PCE with requisite capabilities, e.g., in terms of network layers, algorithms, backup path computation, etc. Two PCE computation models have also been proposed, centralized, and distributed [6]. In the former, all path computation is performed by a single PCE entity. Although this is realistic for intra-domain scenarios, the broader multi-domain case necessitates this PCE to have global resource and policy visibility. Clearly, a single entity also poses notable scalability and reliability limitations.

Meanwhile, distributed computation uses multiple PCE entities to resolve end-to-end paths. This approach also supports "domain-specific" constraints, such as bandwidth/delay in IP domains and wavelength/color continuity in DWDM networks. Moreover, distributed PCE computation is very germane for multi-domain/multi-layer/multi-carrier settings with partial visibility. Herein, two approaches are outlined to handle the contingencies of varying levels of "global" state, i.e., multi-PCE path computation with and without inter-PCE signaling [7]. Overall both of these schemes induce close couplings between inter-domain routing and signaling.

Multi-PCE computation without inter-PCE signaling requires the head-end PCE (i.e., signaled by PCC) to compute a partial or loose route to the destination. This route can be a "domain-level" sequence of border gateways which is then inserted into appropriate downstream signaling messages. Here receiving entities (border gateways) consult their own PCE entities to "expand" the strict-hop sequences across their domains to the next partial/loose hop. This "inter-PCE" approach fits in well with the above-mentioned RSVP-TE LR/ER features. Meanwhile, multi-PCE computation with inter-PCE signaling requires the head-end PCE to recursively request other PCE entities to compute sub-path segments. This setup usually returns a complete strict-hop source-destination route to the requesting PCC and is more suitable for little/no inter-domain visibility. Alternatively, in inter-

carrier settings, this approach may still return a loose route sequence, thereby preserving intra-domain route confidentiality, see [7].

15.3 Research Studies

Many multi-domain research studies have been conducted. However, most of these efforts have focused on single-layer (homogeneous) packet/cell-switching networks, Layers 2–3. Recently some work on multi-domain optical networks (Layer 1) has also emerged. As yet, very few studies have looked at multi-domain/multi-layer inter-networking.

15.3.1 Multi-Domain Packet/Cell-Switching Networks

Many studies have looked at multi-domain ATM (Layer 2) and IP (Layer 3) networks. In particular, graph-based topology abstraction schemes for state reduction [8–11] have been a key focus area. These strategies transform a domain topology into a reduced “virtual” graph with fewer abstract vertices and edges. This is typically done by a designated domain entity, e.g., routing area leader (RAL) [2], which then propagates the abstract link state to other gateway nodes to maintain “globalized” state. Expectedly, the computational entity must have access to intra-domain state. Overall, topology abstraction allows operators to control internal state dissemination and effectively couples intra- and inter-domain routing.

An early study of abstraction in hierarchical ATM PNNI networks is presented in [8]. This effort details algorithms for computing complex node representations to minimize the number of edges/links in a peer group summarization. The findings here show much-reduced computation times and over an order magnitude reduction in domain state. Also, [9] extends abstraction to multi-domain IP QoS networks, summarizing bandwidth state using various topologies, i.e., star, mesh, tree, and spanner graphs. These reductions are tested with various path computation strategies, including widest shortest, shortest-distance, etc. The overall findings confirm improvements in routing scalability and reduced routing fluctuation. Meanwhile, [10] looks at topology aggregation in directed graphs with additive link metrics (delay, cost). This problem is treated from an information-theoretic perspective and related bounds for compression distortion are derived based upon the asymmetry constant. Furthermore, [11] develops schemes to incorporate both bandwidth and delay parameters via line segmentation techniques. Various Dijkstra-based heuristics are then used to compute paths meeting dual bandwidth and delay constraints. Overall, the results show improved success rates and lower crankback messaging loads.

Finally, studies have also looked at QoS enhancements for BGP distance vector routing. For example, [4] proposes messaging extensions to convey

additional QoS state for paths to destination domains/prefixes. This is coupled with pre-engineered QoS-based service level specifications (SLS) between providers along with “meta-QoS class” planes extending across multi-carrier AS. Results for sample Internet topologies show best performance when paths selection is done using both bandwidth and delay constraints. In addition, messaging overheads are shown to be minimal.

15.3.2 Multi-Domain Optical Networks

Multi-domain DWDM networks are also becoming a key focus. To date, most efforts in this area have focused on distributed lightpath RWA and signaling in all-optical and opto-electronic networks. For example, [12] details a domain-by-domain scheme in which gateways maintain complete (alternate) route state across all-optical and opto-electronic networks. Detailed simulation shows the overall effectiveness of the proposed model. However, this setup is more favorable to BGP-type implementations and related resource propagation (path dissemination) results are not shown.

Meanwhile, [13] studies DWDM networks comprising of multiple segments, i.e., domains. Here, three different inter-domain wavelength routing algorithms are developed, i.e., end-to-end (E2E), concatenated shortest path (CSR), and hierarchical routing (HIR). The E2E basically assumes a “flat” globalized graph whereas HIR assumes a hierarchical graph with segments summarized as nodes. Meanwhile CSR simply uses local information to perform segment-by-segment routing. Numerical results with a specialized mesh-torus topology show significant blocking reduction with the E2E scheme for “multi-segment” requests. However, the CSR scheme does not provide any intra-domain state and further inter-domain routing algorithms are not addressed.

Now, very few studies have looked at routing in multi-domain DWDM networks. Here, it is generally very difficult to re-apply IP/MPLS abstraction schemes since resource constraints are different. Hence recent studies have looked at new DWDM-based abstraction schemes. For example, [14] presents a theoretical treatment of abstraction with border node wavelength conversion. Here, various information models are developed and lightpath selection is modeled as a Bayesian decision problem. The findings for bus topologies show that scalable information models can achieve a good trade-off between performance loss and the amount of network state. Nevertheless, inter-domain routing and RWA issues are not addressed.

Meanwhile, [15] tables a hierarchical inter-domain solution using simple-node abstraction. Nevertheless, signaling provisions are not considered here and only the all-optical case is handled. Conversely, [16] presents a much more detailed study of abstraction in multi-domain all-optical/opto-electronic DWDM networks. Namely, graph theoretic RWA schemes (via k-shortest path heuristics) are used to generate full mesh and star abstractions. Furthermore, detailed inter-domain link-state routing protocols and distributed lightpath

RWA algorithms are also specified (using OSPF-TE, RSVP-TE). Findings show very good blocking reduction and lower inter-domain signaling loads.

Dynamic sub-path protection for DWDM networks has also been considered. For example, [17] segments end-to-end lightpaths into independent “protection domains” and implements shared segment protection with/without wavelength conversion. The algorithm pre-defines self-healing cycles in the network topology and attempts to allocate self-healing loops for lightpaths accordingly. Overall, it is shown that maximum sharing can be achieved by properly assigning wavelengths to the shared channels. Nevertheless, these schemes assume full global state, i.e., “flat/single-domain” network, and hence their application in distributed multi-layer settings is not straightforward. Finally, Layer 1 virtual private networks (Layer 1 VPN) provisioning in multi-domain SONET/SDH network has also been studied in [18]. Layer 1 VPN represents a new service model in which transport-layer resources are partitioned to build multiple “virtual infrastructures” over a single physical network. Here, the authors develop a novel OSPF-TE link model to share time-slots between client VPN topologies. Furthermore, non-adjacent L1 VPN clients are interconnected via “virtual link” multi-domain SONET/SDH connections. Overall, results show good sharing efficiency and blocking reductions with shortest-widest path selection. However, all virtual links here are pre-specified (static), as per end-user demands, and inter-domain routing is not done.

15.3.3 Multi-Domain/Multi-Layer Networks

As mentioned in Section 15.1, there is a clear grooming aspect to multi-domain/multi-layer networks. For example, multiple fine-granularity label switched paths (LSP) originating in Layer 3 domains can be aggregated over coarser DWDM lightpaths. In a similar manner, sub-rate SONET/SDH circuits can also be groomed over DWDM lightpaths. Now, even though grooming algorithms have been widely studied over the years, most studies have focused on idealized settings with full resource state knowledge across all layers. The broader topic of distributed multi-domain grooming, i.e., in the presence of limited global state, is largely unexplored and only a handful of studies have been conducted to date.

In particular, [19] presents one of the first studies on distributed SONET-DWDM provisioning in multi-domain networks. Here, the authors define a multi-segment graph model (with boundary grooming) and evaluate various path selection schemes, e.g., centralized (full-knowledge), domain-by-domain (local knowledge), and hierarchical source routing (partial inter-domain knowledge). The latter approach only propagates domain-internal state for a specified granularity levels, although detailed specifications for state compression are not given. Overall, results show much-improved blocking performance with increasing levels of inter-domain state. Furthermore, [20] studies grooming in SONET-DWDM networks and uses threshold-based triggers to add/delete lightpaths. Here, both centralized (global) and distributed schemes

are developed, with the latter only using traversing path state at a node, i.e., no routing. Overall, results show good performances, contingent to grooming ports and request granularities. However, no inter-domain considerations are discussed.

Finally, [21] tables a novel framework for direct Ethernet-over-DWDM integration. This scheme devises an integrated signaling-based approach to simultaneously provision full wavelength and fine granularity Ethernet virtual connections (EVC) at the optical layer. In contrast to conventional provisioning, signaling messages for sub-lambda EVC requests are only processed at the source node of the lightpath. Nevertheless, the authors assume that the underlying DWDM network is relatively static and hence routing is only done at the Ethernet layer. Detailed performance results are presented for various routing update strategies using crankback. Overall, results show minimal increase in setup latencies and signaling loads.

15.4 Open Problems in Multi-domain Grooming

Despite some standards progress, the study of multi-domain/multi-layer grooming remains in its infancy. Now, it is clear that the scope/scale of the problem necessitates distributed, decentralized solutions here. Along these lines, some key open problems need to be addressed, i.e., multi-domain routing, constraint-based path computation/signaling, and survivability.

15.4.1 Multi-layer Routing and State Exchange

Multi-domain routing is a very challenging area, especially in the presence of multiple layers. The most notable issue is how to propagate different types of link or domain state in a such way that a scalable and secure routing framework can be achieved. By and large, multi-domain grooming will benefit from the availability of “global” inter-domain state. Moreover, as per Sections 15.2.3 and 15.3.1, link-state routing is the most suitable scheme for constraint-based provisioning of QoS LSP or TDM/DWDM circuit paths. Hence, it is very desirable to implement some form of hierarchical routing at select border gateways to disseminate various types of link state (as shown in Fig. 15.2):

- **Physical inter-domain links:** These links interconnect nodes in different domains and can span a full range of types, e.g., Gigabit/10 Gigabit Ethernet, SONET/SDH OC-n links, DWDM links, etc. Hence, related multi-layer TEDB must maintain state for all of these link types either via running multiple routing protocol instances or by using a single unified routing protocol. For example, two-level OSPF-TE already provides LSA definitions for both packet-switching links and “non-packet” TDM and DWDM links.

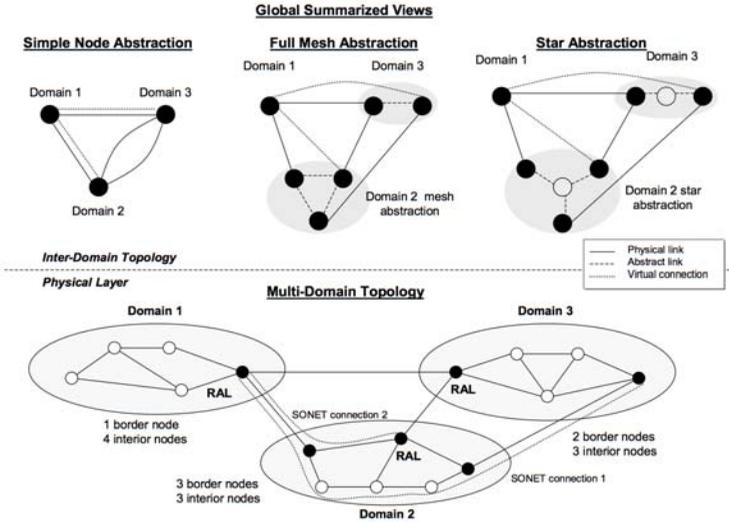


Fig. 15.2. Topology abstraction and link types

- Abstract links:** Abstract links are computed entities used to summarize “layer-specific” domain-level resources and do not reflect physical elements [9]. Although abstraction has been well-studied for IP networks (Section 3.1), circuit-switching networks are more specialized, see initial work in [16]. Hence, further considerations are needed for SONET/SDH grooming links, VCAT features, inverse multiplexing, etc. In addition, new abstractions must capture survivability-related information via abstract links, e.g., fiber/span protection, diversity/risk groups, shared/dedicated resources, etc. In general, expanded attributes can also be developed as these link types need not be constrained to physical link types (in respective domains).
- Virtual link connection:** In general, lower level SONET/SDH and DWDM connections/sub-connections will appear as virtual links to higher-layer devices (i.e., IP/MPLS, Ethernet nodes). These entities can thus be treated as logical TE link (adjacencies) between the two end-points, see Fig. 15.2. Therefore, it will be beneficial to also advertise the resource levels for these entities at the inter-domain level as well, since doing so can yield very good gains in grooming efficiencies. However, there is a clear scalability problem here as the number of such connections can grow in the order of the square of the number of border nodes, i.e., the connection mesh problem as discussed in Section 15.2.2. Hence, a key challenge here will be to develop link aggregation techniques to effectively compress virtual connection link state.

In light of the above, it is also important to consider associated update triggering strategies. To date, various policies have been developed here to achieve update frequency reduction/scalability, e.g., timer-based, absolute change, relative change, and hysteresis-based [20]. All schemes here—except for the former type—are threshold-based. Extensive results for single-domain networks show that relative updates are generally most effective in lowering routing loads and ensuring rapid change sensitivity [20]. Now physical inter-domain link updates can be treated in much the same manner as intra-domain links. Namely, if the underlying resources (bandwidth, timeslots, wavelengths) change as per the chosen metric, an update can be flooded by the sourcing gateway. However, careful analysis is needed to delimit related inter-domain holdoff timers (IHT), which will tend to be larger than corresponding intra-domain values. Virtual link connections can also be treated as such since they are directly associated with underlying connections. Nevertheless, since the number of such connections can be large, further research into scalable triggering policies is needed. Meanwhile, triggering policies for abstract links are more complicated as they pertain to “non-existent” entities. Here, various policies have been developed for abstract IP and DWDM links, see [9, 16], and these can be adapted further. For example, a basic approach is to compute abstractions on a fixed time interval. However, this scheme yields excessive messaging/computational loads for small update intervals and inaccurate state for large update intervals. It is here that relative change measures can give much better performance. For example, [16] uses periodic abstractions but only sends updates if sufficient relative change occurs.

15.4.2 Multi-Layer Grooming and Signaling

Path computation implements the core of the multi-domain grooming function. As opposed to single-domain settings, full topology information in multi-domain networks is generally lacking and path computation has to be done using aggregated (i.e., partial, dated) topology and resource state. Moreover, the existence of multiple layers implies an inherent grooming component, again differing notably from single domain settings which tend to operate at the same layer, e.g., either packet, timeslot, wavelength. Overall, these aspects make multi-layer grooming involving a very challenging problem. In the end, however, the goal here is to resolve and set up full end-to-end routes in accordance with desired constraints and leverage existing standards, particularly GMPLS. Now owing to the distributed nature of the problem it is evident that setup signaling will be required. Although many specific schemes can be envisioned here, few have been analyzed in detail. Hence two broad strategies are outlined here, i.e., hierarchical and per-domain.

Hierarchical Grooming

Hierarchical grooming uses the inter-domain TEDB state to compute “skeleton” routes from the source domain to the destination domain. For

most practical purposes, this path constitutes the end-to-end border node sequence. This information is then coupled with end-to-end signaling mechanisms (RSVP-TE) to expand the full node sequence, i.e., multi-PCE path computation without inter-PCE signaling [7]. Alternatively, this information can also be coupled with PCEP signaling to fill in an exact node sequence, i.e., multi-PCE path computation with inter-PCE signaling [7].

Now consider the actual computation of “skeleton” routes. This is a very complex problem owing to the multiple resource layers involved and the presence of partial/abstracted inter-domain state. Hence, two key areas need detailed investigation here. Foremost, the computing entity (e.g., source domain PCE) must adequately “transform” multi-layer TEDB state to build a “global” multi-layer (granularity) view of the network. The most expedient approach here is to use graph transformations to summarize all link types—physical, abstract, and virtual (Section 15.4.1). This topic is largely unexplored, although augmented graph models have been developed in [9] for multi-layer SONET-DWDM grooming networks (full global state assumption). Clearly, new schemes are needed for adding edges/vertices to capture virtual link connections as well. Next, modified shortest-path schemes need to be studied to compute constrained “skeleton” paths over these transformed graphs. It is here that the grooming aspect can be considered, particularly when “higher-layer” sub-rate connections are to be routed over lower-layer circuits (virtual link connections).

Most regular (non-multi-domain) grooming studies use a “two-step” computation approach. Namely, setup is first attempted over existing higher-layer “virtual” links and pending failure, re-attempted at lower transport layers. Overall, this approach yields very good resource efficiency/packing and is equally relevant in inter-domain scenarios. For example, the initial attempt can search for a “skeleton” route on only those links in the augmented graph (physical inter-domain, abstract, virtual connection) which run at the same granularity/layer as the sourcing node. Here various graph theoretic schemes can be developed to achieve trade-offs between objectives such as hop count, load balancing, minimum cost, e.g., see studies on widest-shortest, shortest-widest, and minimum cost [9].

If a feasible “skeleton” path is not found, further attempts can try to set up new underlying virtual link connection between one/more border gateways. However, this step opens up a whole new dimension that is not present in single-layer multi-domain settings. Namely, the main challenge is to select the actual border node pair(s) between which to initiate virtual connection link(s). This concern has not been addressed in most grooming studies which exclusively assume that all higher layer nodes are directly connected to underlying transport nodes. Clearly this is not the case in general settings in which interior IP/MPLS nodes will not be directly connected to lower-layer SONET/SDH or DWDM transport nodes. One possibility is to use exhaustive search algorithms to enumerate multiple virtual links combinations, e.g., as

done in [16]. Another simpler possibility may be to pre-engineer static virtual connection links using offline optimizations.

Upon computation of a feasible “skeleton” path, appropriate setup signaling schemes are needed as well. In general, these schemes can leverage the ubiquitous RSVP-TE protocol. Specifically, resources can be reserved on physical inter-domain and “tunneled” virtual connection links and path sequences “expanded” for loose route segments. With regards to the latter, border gateways receiving a setup request will basically run intra-domain algorithms to route requests across their domains. Now if the incoming request is of the same granularity, the border node can compute (or query PCE to compute) a regular traversing as per path constraints. However, in the more general case where the border node represents a lower layer, further parameter translation and intra-domain grooming will be needed, and a wide range of existing (intra-domain) grooming schemes can be re-applied. Furthermore, new RSVP-TE proposals are also tabling LSP “stitching” features to inter-connect diverse end-to-end “segments” (layer paths) into a single connection. Again, these can be leveraged as well.

Overall, hierarchical computation can achieve some level of path “optimality” across domains, i.e., contingent to level of inter-domain state. An “optimal” path here is defined as that computed in the absence of domain partitioning, i.e., for “flat” topology [7]. However, detailed studies are really needed to assess the impact of abstracted state on such “optimality”.

Per-Domain Grooming

Per-domain distributed grooming is ideal for settings with minimal inter-domain visibility. Namely, end-to-end routes are expanded in a domain-by-domain manner with each domain specifying its “next-hop” along with the complete internal route via an egress gateway. Here, intra-domain path expansion will be largely similar to the above-detailed hierarchical approach. However at the inter-domain level, the main challenge is how to specify the next domain. If no inter-domain state is available, the only solution may be to make fixed choices, i.e., pre-specified or by consulting BGP tables [7]. Alternatively, it may be much more beneficial to incorporate “global” state to dynamically the downstream domains, if such state is available. Nevertheless, the detailed evaluation of these schemes in mixed-layer networks has not been conducted. In general, per-domain path computation (without inter-domain state) will suffer from higher blocking and lower resource efficiencies as compared to hierarchical computation. Hence, resultant routes will likely be feasible and opposed to “optimal”.

Note that per-domain grooming performance can be improved by leveraging inter-domain crankback signaling. Although [22] studies this approach in SONET-DWDM networks, more investigations are needed. Specifically, commensurate schemes can be designed using intermediate domain or full end-to-end crankback strategies. Furthermore, it may be beneficial to also include active path state in crankback messages in order to improve the

re-try process. Although multi-layer crankback can yield improved resource efficiencies, it poses key trade-offs in inter-layer control plane complexity. Namely, this solution may lower inter-layer routing overheads (e.g., versus hierarchical computation) but at the same time will likely increase inter-layer signaling and setup latencies. These issues need to be closely studied in order to determine the most amenable grooming strategies, i.e., as per connection granularities, arrival rates, holding times, etc.

15.4.3 Survivability

With limited or no visibility into domain-level state, setting up survivable connections is an extremely complicated issue. Hence, the simple approach of computing diverse routes by source nodes (with partial global state) can only guarantee diversity at the physical inter-domain link level. Clearly, further provisions are needed to avoid link overlap during subsequent intra-domain explicit route expansion even with disjoint abstract links computed. Overall, survivable grooming is a relatively new topic area. For example, [23] has recently proposed some protection schemes for SONET-DWDM grooming networks, e.g., protection-at-lightpath (PAL) and protection-at-connection (PAC). Nevertheless, extending these concepts into multi-domain networks lacking full resource diversity state is very challenging. Furthermore, it is also very desirable to extend some form of multi-tiered survivability support as well, e.g., dedicated, shared, non-protected, etc. To address this topic, two survivability strategies can be considered, protection and restoration.

Protection schemes pre-compute diverse primary/backup routes and are well suited for stringent services. Now in the context of multi-domain/multi-layer networks, a simple approach can be to apply SONET strategies in which dual/multi-homed inter-domain links are coupled with robust intra-domain protection schemes. However, these setups are very costly and resource inefficient. Hence, more elaborate end-to-end domain disjoint techniques can be devised, which are generally more amenable to hierarchical grooming with inter-domain state visibility. For example, loose routing algorithms can be extended to compute domain disjoint primary/back “skeleton” paths over augmented inter-layer graphs. Associated setup (loose route expansion) sequences can then be signaled in parallel. Alternatively, the domain-disjoint requirement can be relaxed to allow sharing of abstract links between primary/backup routes. However, the latter approach mandates complicated signaling provisions to ensure intra-domain disjointness in common domains. Again, many of these concerns are open issues and the combined impact of intra/inter-domain schemes on end-to-end recovery still needs to be addressed.

Conversely protection can also be incorporated with the per-domain grooming approach of Section 15.4.2. For example, primary/backup paths can be signaled in a two-step sequential manner where the working route is used to avoid backup link overlaps. Obviously, this approach will have higher setup delays and will be susceptible to “trap” topologies [7]. To address these

concerns, modified crankback schemes can be considered to achieve protection path diversity with parallel signaling, e.g., as tabled in RSVP-TE extensions for record, exclude and associated route objects (RRO, XRO, ARO) [24]. To date, however, most of these concepts have only appeared various standards draft submissions (MPLS-level only), and more detailed analyses and refinements are indeed lacking.

Finally, multi-domain restoration can also be done using post-fault crank-back recovery and this is most germane for connections with lower latency/guarantee stringencies. Along these lines, various signaling provisions need to be devised, e.g., such as proposed RSVP-TE XRO extensions. However, further studies are needed to analyze the trade-offs between resource efficiencies and signaling loads and also compare versus pre-provisioned protection.

15.5 Conclusion

This chapter studies multi-domain grooming in heterogeneous layered networks, a challenging and largely unexplored problem. Along these lines related standards and research studies are surveyed and key open research problems identified. In particular, the latter issues pertain to multi-layer/multi-domain routing, path computation and signaling, and survivability. Finally, future studies can delve into even more advanced multi-domain problems, e.g., VPN provisioning, advance reservation, etc.

References

1. N. Ghani, et al., "Metropolitan Optical Networks," *Optical Fiber Telecommunications IV*, Academic Press, pp. 329–403, March 2002.
2. G. Bernstein, B. Rajagopalan, D. Saha, *Optical Network Control-Architecture, Protocols and Standards*, Addison Wesley, Boston 2003.
3. N. Rao, et al., "Ultra Science Net: Network Testbed for Large-Scale Science Applications," *IEEE Communications Magazine*, Vol. 3, No. 4, pp. S12–S17, November 2005.
4. D. Griffin, et al., "Interdomain Routing Through QoS-Class Planes," *IEEE Communications Magazine*, Vol. 45, No. 2, pp. 88–95, February 2007.
5. T. Lehman, et al., "Control Plane Architecture and Design Considerations for Multi-Service, Multi-Layer, Multi-Domain Hybrid Networks," *IEEE INFOCOM 2007 HSN Workshop*, Anchorage, Alaska, May 2007.
6. A. Farrell, "A Framework for Inter-Domain Traffic Engineering," *IETF Re-quest for Comments (RFC) 4726*, November 2006.
7. A. Farrell, J. Vasseur, J. Ash, "A Path Computation Element (PCE)-Based Architecture," *IETF Request for Comments (RFC) 4655*, August 2006.
8. I. Iliadis, "Optimal PNNI Complex Node Representations for Restrictive Costs and Minimal Path Computation Time," *IEEE/ACM Transactions on Network-ing*, Vol. 8, No. 4, August 2000.

9. F. Hao, E. Zegura, "On Scalable QoS Routing: Performance Evaluation of Topology Aggregation," *IEEE INFOCOM 2000*, pp. 147–156, 2000.
10. B. Awerbuch, Y. Shavitt, "Topology Aggregation for Directed Graphs," *IEEE/ACM Trans. on Networking*, Vol. 9, No. 2, pp. 82–90, February 2001.
11. K. Liu, K. Nahrstedt, S. Chen, "Routing with Topology Abstraction in Delay-Bandwidth Sensitive Networks," *IEEE/ACM Transactions on Networking*, Vol. 12, No. 1, pp. 17–29, February 2004.
12. X. Yang, B. Ramamurthy, "Inter-Domain Dynamic Routing in Multi-Layer Optical Transport Networks," *IEEE GLOBECOM 2003*, San Francisco, CA, December 2003.
13. Y. Zhu, A. Jukan, M. Ammar, "Multi-Segment Wavelength Routing in Large-Scale Optical Networks," *IEEE ICC 2003*, Anchorage, AL, May 2003.
14. G. Liu, C. Ji, V. W. S. Chan, "On the Scalability of Network Management In-formation for Inter-Domain Light-Path Assessment," *IEEE/ACM Transactions on Networking*, Vol. 13, No. 1, pp. 160–172, January 2005.
15. S. Sanchez-Lopez, et al., "A Hierarchical Routing Approach for GMPLS-Based Control Plane for ASON," *IEEE ICC 2005*, Seoul, Korea, June 2005.
16. Q. Liu, et al., "Hierarchical Inter-Domain Routing and Lightpath Provisioning in Optical Networks," *OSA Journal of Optical Networking*, Vol. 5, No. 10, pp. 764–774, October 2006.
17. P. Ho, H. T. Mouftah, "A Novel Survivable Routing Algorithm for Segment Shared Protection in Mesh WDM Networks with Partial Wavelength Conversion," *IEEE JSAC*, Vol. 22, No. 8, pp. 1539–1548, October 2004.
18. N. Ghani, et al., "Layer 1 VPN Services in Distributed Next-Generation SONET/SDH Networks With Inverse Multiplexing," *OSA Journal of Optical Networking*, Vol. 5, No. 5, pp. 367–382, May 2006.
19. Y. Zhu, et al., "End-to-End Service Provisioning in Multi-Granularity Multi-Domain Optical Networks," *IEEE ICC 2004*, Paris, France, June 2004.
20. R. Alnuweri, et al., "Performance of New Link State Advertisement Mechanisms in Routing Protocols with Traffic Engineering Extensions," *IEEE Comm. Magazine*, Vol. 42, No. 5, pp. 151–162, May 2004.
21. A. Hadjiantonis, et al., "Evolution to a Converged Layer 1, 2 in a Global-Scale, Native Ethernet Over WDM-Based Optical Networking Architecture," to appear in *IEEE Journal of Selected Areas in Communications*, 2007.
22. I. Widjaja, et al., "A New Approach for Automatic Grooming of SONET Circuits To Optical Express Links," *IEEE International Conference on Communications (ICC) 2003*, Anchorage, AL, May 2003.
23. C. Ou, et al., "Traffic Grooming for Survivable WDM Networks-Shared Protection," *IEEE JSAC*, Vol. 21, No. 9, pp. 1367–1383, November 2003.
24. C. Lee, et al., "Exclude Routes-Extensions to RSVP-TE," *IEEE Communications Magazine*, November 2006.

Grooming of Scheduled Demands in Multi-Layer Optical Networks

Maurice Gagnaire, Josué Kuri*, Elias A. Doumith

16.1 Introduction

Introducing multiple switching granularities in transport networks aims in part at reducing their cost by taking advantage of traffic aggregation. As a matter of fact, an appropriate aggregation of low-rate connections into high-rate connections¹ and the use for the latter of a switching technology whose cost-per-bit is lower than for the former can result in lower infrastructure costs. The economic gain, when compared to single switching granularity networks, depends on both traffic-related factors such as the time/space distribution of connections and equipment or network architecture factors such as the cost of ports and matrices for each switching granularity, the topology of the network, the provisioning algorithms used for routing and grooming of connections, etc.

In other chapters of this book, traffic grooming has been investigated in contexts such as protection/restoration, multicast, and multi-domain networks. In this chapter, we consider the problem of efficiently grooming (a) electrical “subwavelength” demands into lightpaths, and (b) lightpaths into wavebands. Whereas most investigations deal with either permanent or random traffic demands, we focus in this chapter on a particular class of traffic demands referred to as scheduled demands, or SxDs, introduced in [1, 2]. A scheduled traffic demand corresponds to a connection or set of connections for which the network operator knows in advance the set-up and tear-down dates. These connections could be requested, for example, by a company to connect its headquarters to its production centers during office hours, or to interconnect its data centers during the night, when database backups are performed. Since companies tend to have well-established operations and procedures, they can provide relatively accurate information to the network operator

* This work was carried out while the author was at ENST Paris. The opinions expressed in this paper are those of the author and not necessarily those of Orange Labs.

¹ For example, Low Order (LO) and High Order (HO) SONET/SDH connections.

about when the bandwidth is really needed. An operator can leverage the knowledge about the timing of demands to make a more effective use of its network resources.² For example, a same set of switch ports and transponders can be allocated to multiple connections routed over a same path if the operator knows in advance that the connections are time-disjoint. Conceptually, a permanent connection is a special case of a scheduled demand for which the set-up and tear-down dates are $-\infty$ and ∞ , respectively. In practice, the set-up and tear-down dates of a permanent connection correspond to the beginning and end dates of the planning period considered in a network design exercise. Therefore, by solving the problem of provisioning scheduled demands, we also solve the problem of provisioning permanent demands.

From a mathematical modeling point of view, an appealing characteristic of scheduled demands is that they capture the dynamic changes of the traffic load in a *deterministic* manner, which eases the use of well-known optimisation techniques such as heuristics [3, 4, 5], meta-heuristics [6], linear programming [7, 8, 9], etc.

A scheduled demand is formally characterized by a tuple (s, d, n, α, β) in which s and d represent the source and destination nodes, n is the requested capacity, and α and β represent the set-up and tear-down dates. Two forms of scheduled demands are considered in this chapter: Scheduled Lightpath Demands (SLDs) and Scheduled Electrical Demands (SEDs), also known as “subwavelength” demands. For SLDs, the capacity n is expressed in number of lightpaths (each one with a nominal rate of either 2.5 or 10 Gbps), and for SEDs, n is expressed as a fraction of a lightpath’s capacity. For example, $n = 0.4$ corresponds to 1 Gbps with 2.5 Gbps lightpaths.

For the problem of grooming SEDs into SLDs, we develop in this chapter a mathematical formulation to quantify the number of Electrical Cross-Connect (EXC) and Wavelength Cross-Connect (WXC) ports required to provision a set of SEDs in a given topology [10, 11, 12, 13]. For the problem of grooming SLDs into wavebands, we develop a different formulation which quantifies the monetary cost of provisioning a set of SLDs in a network in which an integrated Wavelength Cross-Connect (WXC) / Waveband Cross-Connect (BXC) equipment is present in each node.³ We refer to the former as the electrical grooming problem and to the latter as the optical grooming problem.

² This is analogous to what airlines do: passengers that buy tickets well in advance are rewarded with lower fares because early bookings reduce the uncertainty of the future demand, which allows the airline to make a more effective use of its resources.

³ An EXC is essentially the same equipment that has been referred to as a DXC in earlier chapters, and WXC and BXC are both OXCs, operating at different wavelength selectivities. - Editors

16.2 Architecture of Network Nodes

Figure 16.1 depicts the architecture of a multi-granularity switching cross-connect (MG-OXC) that integrates an electrical cross-connect (EXC), a wavelength cross-connect (WXC), and a waveband cross-connect (BXC). The composite signal received on the input fibers of the BXC is demultiplexed to extract individual wavelengths which are in turn multiplexed into wavebands that arrive at the input ports of the BXC switching fabric. The BXC cross-connects input wavebands to either output or drop ports. Dropped wavebands are demultiplexed to extract individual wavelengths which are cross-connected by the WXC to either output or drop ports. Dropped wavelength connections are terminated by the EXC, which extracts “subwavelength” connections (or flows) that are sent to client equipment (e.g., routers) through the EXC drop ports. The MG-OXC has the capability of dropping full lightpaths directly from the WXC to client equipment. In the opposite direction, “subwavelength” connections (or flows) received from client equipment through the EXC add ports are groomed into lightpaths, which are received by the WXC through its add ports and groomed into wavebands that are received by the BXC through its add ports and groomed into fibers. The WXC of the MG-OXC has the capability of adding lightpaths directly from client equipment.

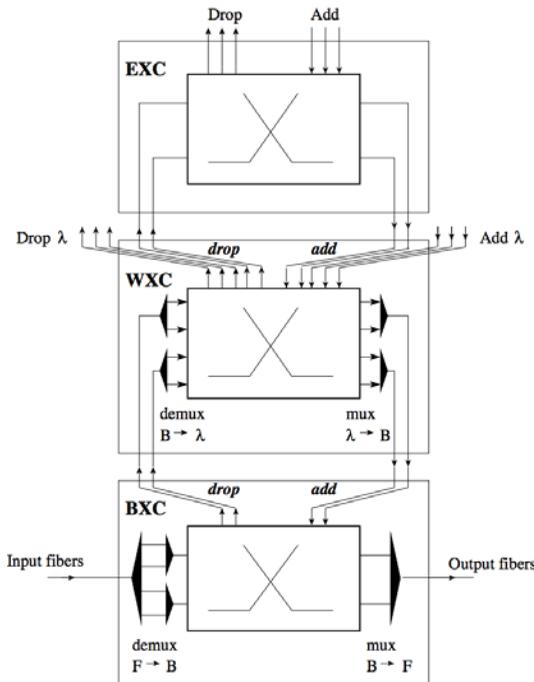


Fig. 16.1. Multilayer node architecture

In this chapter, we deal with electrical and optical grooming as two separate problems. In the former case, the network consists of EXC/WXC nodes (no BXC), whereas in the latter, the network consists of WXC/BXC nodes (no EXCs).

16.3 Rationale of Grooming Strategies

The goal of grooming strategies in multi-layer networks is to reduce the overall management burden and network cost (compared to single-layer networks) by efficiently grouping a large number of low-rate connections into a relatively small number of high-rate connections. The overall network cost can be reduced if the cost-per-bit is lower in the network layer bearing the high-rate connections than in the upper layers, and if the cost of grouping/ungrouping connections does not outweigh the savings of the high-rate connections' layer. Architectures like the one of Fig. 16.1 may reduce the network cost if BXC are less expensive than WXC, and WXC are in turn less expensive than EXC. This tends to be the case with state-of-the-art technology since EXCs require electronic switching fabrics and opto-electronic converters on I/O ports, whereas WXC and BXC require potentially less opto-electronic devices and can be based on less expensive optical switching fabrics.

The presence of transit traffic in a multi-granularity switching cross-connect can reduce the number of required EXC and WXC ports since wavebands bearing transit lightpaths are switched directly from input to output ports of the BXC without requiring any port on the WXC. Likewise, lightpaths bearing transit "sub-wavelength" traffic are switched directly from input to output ports of the WXC without requiring any port on the EXC. Lightpaths on the input ports of the WXC are either recombined into wavebands through the output ports or sent to either client equipment or input ports of the EXC. The EXC extracts "sub-wavelength" connections (or flows) which are sent to client equipment through the drop ports or recombined into new lightpaths sent to the WXC through the output ports.

Provisioning in a multi-layer network involves essentially the definition of both, a path for each SEDs or SLDs, and the way the demands are grouped and ungrouped (potentially several times) along the path. In a single-layer network, a simple strategy consisting in routing connections along the shortest path in terms of hops results in the lowest cost since the total number of required cross-connect ports is minimized. In a multi-layer network, the provisioning aimed at minimizing cost is less obvious since demands can be deviated from their shortest paths in order to create long, efficiently utilized high-rate connections to reduce the overall network cost. The provisioning process becomes even more complex when the demands are SxDs, instead of permanent connections, since the time-disjointness among demands needs to be taken into account in order to reuse network resources as much as possible.

16.3.1 Electrical Grooming

The characteristic of the SEDs lies in the fact that their flow is smaller than the capacity of a lightpath. This particularity is taken into account by use of electrical aggregation. The grooming of multiple SEDs consists in multiplexing their electrical flows onto the same Grooming Lightpath (GL) in an EXC. Grooming requires all the considered SEDs to share at least one common fiber link and all of them to be active during a common period of time. In order to explain the rationale of electrical traffic grooming, let us consider the simple case of the aggregation of two SEDs into a GL. These two SEDs, referred to as δ_1 and δ_2 , are characterized by the tuples $(s_1, d_1, n_1, \alpha_1, \beta_1)$ and $(s_2, d_2, n_2, \alpha_2, \beta_2)$, respectively.

Figure 16.2 illustrates the network topology that supports δ_1 and δ_2 . Below the illustration of the network we have plotted the time diagram of δ_1 and δ_2 . Both requests are active from α_2 to β_1 and they have a common path between node g_1 and node g_2 .

When the network does not have grooming capabilities, each SED is routed using a separate GL between its source and its destination nodes, as illustrated in Fig. 16.3. In other words, δ_1 is routed using the GL φ_1 between s_1 and d_1 , while δ_2 is routed using the GL φ_2 between s_2 and d_2 . As a result, two optical channels are needed on each fiber link between nodes g_1 and g_2 .

On the other hand, when the network has grooming capabilities, δ_1 and δ_2 can be groomed together between nodes g_1 and g_2 if $n_1 + n_2 \leq C_\omega$ as illustrated in Fig. 16.4 (C_ω is the nominal capacity of a lightpath). This grouping creates an aggregated demand $\delta_a = (g_1, g_2, n_1 + n_2, \alpha_2, \beta_1)$ borne by GL φ_a . However, since δ_1 is also active during period $[\alpha_1, \alpha_2]$ and δ_2 is also active during period $[\beta_1, \beta_2]$, two additional requests must be considered: $\delta_f = (g_1, g_2, n_1, \alpha_1, \alpha_2)$ to take into account the remaining active period of δ_1 and $\delta_g = (g_1, g_2, n_2, \beta_1, \beta_2)$

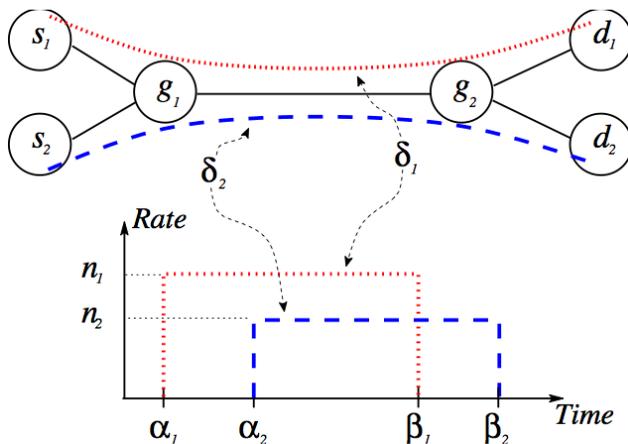


Fig. 16.2. Time diagram of two SEDs

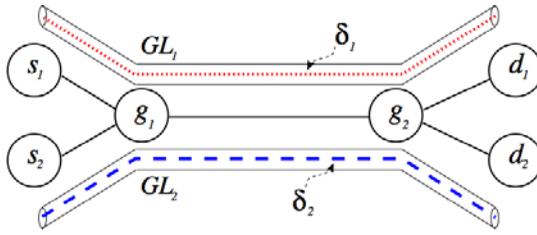


Fig. 16.3. Network without grooming capabilities

to take into account the remaining active period of δ_2 . In addition, we still need to take into account the non-common segments of the paths of δ_1 and δ_2 . This can be done by introducing four additional traffic requests δ_b , δ_c , δ_d , and δ_e :

- Request $\delta_b = (s_1, g_1, n_1, \alpha_1, \beta_1)$ carried by GL φ_b represents a segment of the path of δ_1 totally disjoint from the path of δ_2 . It connects s_1 , the source node of δ_1 to the grooming node g_1 .
- Request $\delta_c = (g_2, d_1, n_1, \alpha_1, \beta_1)$ carried by GL φ_c represents another segment of the path of δ_1 totally disjoint from the path of δ_2 . It connects the grooming node g_2 to d_1 , the destination node of δ_1 .

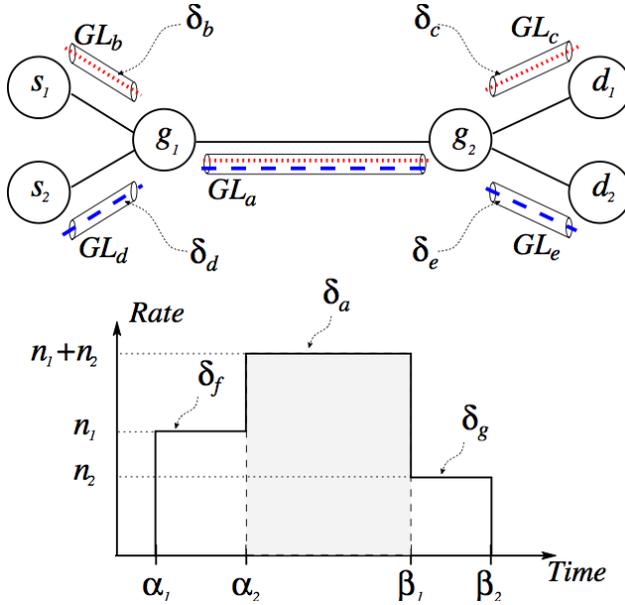


Fig. 16.4. Network with grooming capabilities

- Request $\delta_d = (s_2, g_1, n_2, \alpha_2, \beta_2)$ carried by GL φ_d represents a segment of the path of δ_2 totally disjoint from the path of δ_1 . It connects s_2 , the source node of δ_2 , to the grooming node g_1 .
- Request $\delta_e = (g_2, d_2, n_2, \alpha_2, \beta_2)$ represents another segment of the path of δ_2 totally disjoint from the path of δ_1 . It connects the grooming node g_2 to d_2 , the destination node of δ_2 .

To sum up, grooming together requests δ_1 and δ_2 creates the aggregated request δ_a and a set of six additional requests. Some of these requests ($\delta_b, \delta_c, \delta_d$, and δ_e) take into account the non-common segments of the paths of δ_1 and δ_2 . The remaining requests (δ_f and δ_g) take into account the non-common periods of time. Depending on the characteristics of the initial requests δ_1 and δ_2 , some of these additional requests may not exist. However, those which do exist form the set of marginal demands. As a result of the grooming process, when grooming two SEDs, the set of all the SEDs must be modified by adding the aggregated request δ_a and the set of marginal demands and by removing the original requests δ_1 and δ_2 .

At node g_1 , the two demands δ_1 and δ_2 must be switched from the WXC to the EXC in order to be groomed together. Hence two receiving r_3 electrical ports and two emitting o_3 ports are needed at this node. Once groomed together onto the the GL φ_a , the resulting grooming lightpath is switched back to the WXC using an emitting e_3 electrical port and a receiving o_3 optical port. Arriving at node g_2 , the GL φ_a is switched from the WXC to the EXC using an emitting o_3 optical port and a receiving r_3 electrical port. At the EXC of node g_2 , the initial requests δ_1 and δ_2 are rebuilt and each one is switched to its destination using two emitting e_3 electrical ports and two receiving o_3 optical ports. In brief, in order to groom these two requests, six additional electrical ports and six additional optical ports are needed. Figure 16.5 details the ports involved in the grouping and ungrouping operations at nodes g_1 and g_2 , respectively.

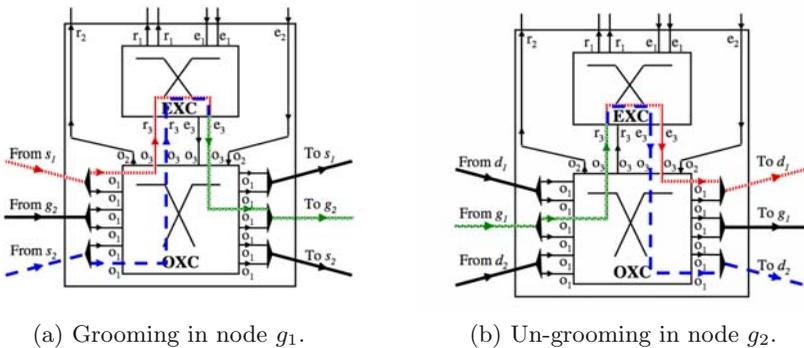


Fig. 16.5. Grooming operations in nodes g_1 and g_2

16.3.2 Optical Grooming

For the optical grooming problem, we consider a network of WXC/BXC nodes (no EXCs). We assume that WXCs provide full wavelength conversion. This simplifies wavelength assignment since the wavelength continuity constraint does not exist. As explained in Section 16.2, the lightpaths at the output (input) ports of the WXC are multiplexed (demultiplexed) into (from) wavebands which are directly added in (dropped from) a BXC. Thus, the WXCs are not directly connected to each other but through waveband-switching connections between BXCs. This means that lightpaths are instantiated over a logical topology of waveband-switching connections. These connections cannot be directly added in or dropped from the BXC. The number of WXC I/O ports is a multiple of the waveband size.

Figure 16.6 shows how three lightpaths can be instantiated over a logical topology formed by two waveband-switching connections, one between nodes 1 and 3 and another between nodes 3 and 4. Two lightpaths are added in the WXC of node 1 using the WXC add ports. They are then multiplexed into a waveband-switching connection, which is added into the BXC of the same node using one BXC add port. The resulting waveband-switching connection is multiplexed into a fiber. In node 2, the connection is demultiplexed from the incoming fiber, switched to an output port and multiplexed into a fiber. In node 3, the connection is demultiplexed from the incoming fiber and dropped using one BXC drop port. This connection is demultiplexed in order to retrieve the two lightpaths which are then switched from the input to the output ports of the node. The lightpaths are multiplexed together with a third lightpath added in node 3 into a waveband-switching connection which is added in the BXC of the node. The connection is multiplexed into a fiber. In node 4, the connection is demultiplexed from the incoming fiber and dropped. The three

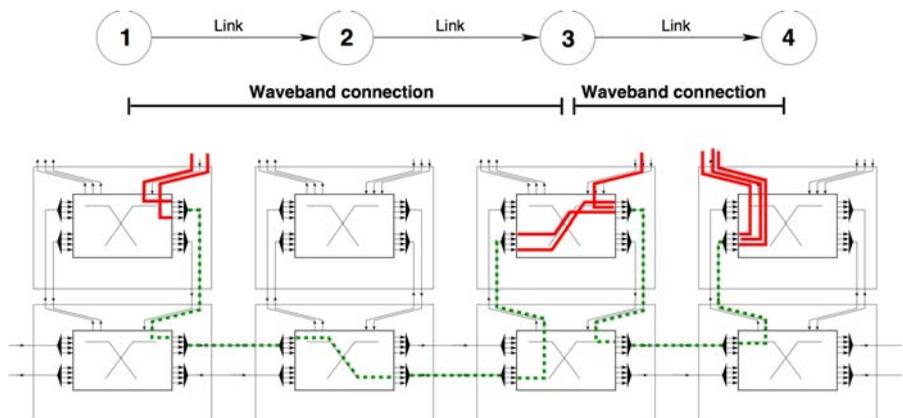


Fig. 16.6. A possible configuration of MG-OXCs used to set up three lightpaths

lightpaths are demultiplexed from the waveband-switching connection and dropped using the WXC drop ports of the node.

Waveband-switching introduces an intermediate network layer between the physical network and the SLDs. A logical topology of waveband-switching connections must be defined and mapped on the physical network. An SLD is instantiated by defining a path on the logical topology and assigning waveband-switching connections on each arc of the path to the SLD's lightpaths. In this context, grooming refers to the aggregation (and disaggregation) of lightpaths into waveband-switching connections of the logical topology. Thus, the optical grooming problem under consideration consists in defining a logical topology of waveband-switching connections, routing these connections on a physical network, routing a set of SLDs on the logical topology, and assigning the waveband-switching connections of the logical topology to the SLDs such that the network cost is minimal. The cost of the network is equal to the sum of the costs of the MG-OXC required to implement the solution. The problem is schematically illustrated by Fig. 16.7. An instance of the problem is defined by a set of SLDs and the topology of a physical network. The MG-OXC cost function models the cost of a MG-OXC as a function of the number of WXC and BXC ports of the MG-OXC (the function is defined in Equation 16.16). The network functional model defines the architecture of the switches, including the definition of the size of the wavebands (the number of lightpaths that can be groomed into a waveband-switching connection). The objective function defines the optimality criterion to be satisfied, for example, the minimization of the network cost. The solution to the problem consists of a set of Routed Scheduled Band Groups (RSBGs) and of an assignment of SLDs to RSBGs. An RSBG is similar to an SLD in that it is defined by a tuple $(s, d, n, \alpha, \beta, P)$ where s and d are the source and destination BXC of the RSBG, n is the number of waveband-switching connections in the group, α and β are the set-up and tear-down dates of the RSBG, and P is the route in the physical network. The set-up date α (tear-down date β) of an RSBG is the earliest set-up (latest tear-down) date of any of the SLDs assigned to this RSBG.

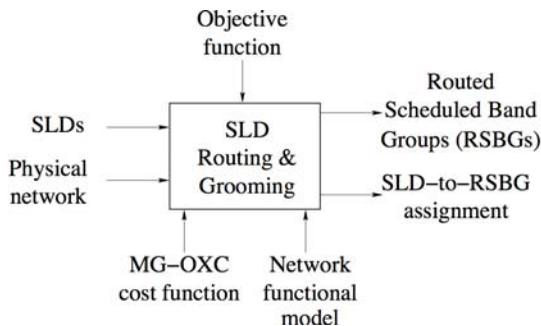


Fig. 16.7. Schematic representation of the optical grooming problem

16.4 Mathematical Formulation

In this section, we develop the mathematical formulations used to quantify the network cost in the electrical and optical grooming problems. The notations in Sections 16.4.1 and 16.4.2 are independent, that is, the symbols used in one subsection are not related to the symbols of the other.

16.4.1 Electrical Grooming

The design of a grooming-capable network is formally stated below. The inputs to the problem are:

- A physical topology represented by an arc-weighted symmetrical directed graph $G = (V; E; w)$. $V = \{v_1, v_2, \dots, v_N\}$ is the set of network nodes and $E = \{e_1, e_2, \dots, e_F\}$ is the set of physical links interconnecting the nodes. Nodes correspond to EXC/WXC multi-layer switches and links correspond to the fibers of the network. Though links are unidirectional, we assume that there are equal number of fibers linking two nodes in different directions. Links are assigned weights $w : E \mapsto \mathbb{R}^+$ which may correspond to the physical length or cost of the links. $N = |V|$ and $F = |E|$ represent the number of nodes and links in the network, respectively.
- A set $\mathcal{D} = \{\delta_1, \delta_2, \dots, \delta_M\}$ of M SxD requests to be routed on the network. These requests can be divided into a set $\mathcal{D}^L = \{\delta_1^L, \delta_2^L, \dots, \delta_{M^L}^L\}$ of M^L SLD requests and a set $\mathcal{D}^E = \{\delta_1^E, \delta_2^E, \dots, \delta_{M^E}^E\}$ of M^E SED requests. SLDs (δ_i^L) and SEDs (δ_i^E) are represented by a tuple $(s_i, d_i, n_i, \alpha_i, \beta_i)$ where $s_i, d_i \in V$ are the source and destination nodes of the demand, n_i is the traffic rate requested by the connection, and α_i and β_i are its setup and tear-down dates. For SLDs, n_i is an integer, while for SEDs n_i is a real number smaller than 1.
- A set of available routes $\mathcal{P}_i = \{P_{i,1}, P_{i,2}, \dots, P_{i,K}\}$ connecting the source node s_i to the destination node d_i of request δ_i^L or δ_i^E . For each request, we compute beforehand K alternate shortest paths in terms of effective length connecting its source node to its destination node using the algorithm described in [14] (if as many paths exist, otherwise we consider the available ones).

The set-up and tear-down dates of all these requests are to be taken into account because they represent time instants when a change in the traffic flow is observed. Let \mathcal{E} be the ordered set grouping the set-up and tear-down dates of all the requests in the set \mathcal{D} . Because some requests may have the same set-up and tear-down dates, the number $\mathbb{E} = |\mathcal{E}|$ of time instants in \mathcal{E} is less than or equal to $2 \times M$.

$$\begin{aligned} \mathcal{E} &= \left(\bigcup_{i=1}^M \alpha_i \right) \cup \left(\bigcup_{i=1}^M \beta_i \right) \\ &= \{\epsilon_1, \epsilon_2, \dots, \epsilon_{\mathbb{E}}\} \text{ such as } \epsilon_1 < \epsilon_2 < \dots < \epsilon_{\mathbb{E}} \end{aligned} \quad (16.1)$$

In the case of SEDs, a physical path $P_{i,\rho_i} \in \mathcal{P}_i$ is assigned in a first step to the request δ_i^E ($1 \leq i \leq M^E$). P_{i,ρ_i} is the ρ_i^{th} shortest path between s_i and d_i ($1 \leq \rho_i \leq K$). Once the physical path P_{i,ρ_i} is chosen, we determine in a second step the intermediate nodes of P_{i,ρ_i} in which δ_i^E must be switched to the electrical level for grooming. In this way, we obtain the set of grooming lightpaths that the request traverses in the logical topology.

A physical path P_{i,ρ_i} , composed of ℓ_{i,ρ_i} hops, has $(\ell_{i,\rho_i} - 1)$ intermediate nodes and can be decomposed into $\mathbb{L}_{i,\rho_i} = 2^{\ell_{i,\rho_i} - 1}$ different sets of grooming lightpaths. For example, the physical path $P_{i,\rho_i} = 1 - 2 - 3 - 4$ between nodes 1 and 4 can be decomposed into the following $2^{3-1} = 4$ different sets of grooming lightpaths:

- $L_{i,\rho_i,1} = \{[1 - 2 - 3 - 4]\}$
- $L_{i,\rho_i,2} = \{[1 - 2], [2 - 3 - 4]\}$
- $L_{i,\rho_i,3} = \{[1 - 2 - 3], [3 - 4]\}$
- $L_{i,\rho_i,4} = \{[1 - 2], [2 - 3], [3 - 4]\}$

$L_{i,\rho_i,\lambda_i,\rho_i}$ is the λ_i^{th} possible decomposition of the physical path P_{i,ρ_i} into a set of grooming lightpaths ($1 \leq \lambda_i,\rho_i \leq \mathbb{L}_{i,\rho_i}$). Let $\mathcal{L}_{i,\rho_i} = \{L_{i,\rho_i,1}, L_{i,\rho_i,2}, \dots, L_{i,\rho_i,\mathbb{L}_{i,\rho_i}}\}$ be the set of all possible decompositions of the path P_{i,ρ_i} into its inherent sets of grooming lightpaths.

The solution of the optimization problem consists in assigning to the request δ_i^E a physical path $P_{i,\rho_i} \in \mathcal{P}_i$ and a set of grooming lightpaths $L_{i,\rho_i,\lambda_i,\rho_i} \in \mathcal{L}_{i,\rho_i}$. This solution is represented by $\Pi_{\mathcal{D}^E,\rho,\lambda} = \{(P_{1,\rho_1}, L_{1,\rho_1,\lambda_{1,\rho_1}}), (P_{2,\rho_2}, L_{2,\rho_2,\lambda_{2,\rho_2}}), \dots, (P_{M^E,\rho_{M^E}}, L_{M^E,\rho_{M^E},\lambda_{M^E,\rho_{M^E}}})\}$.

Our objective is to satisfy all the requests at the lowest cost. The network cost is expressed in terms of the number of electrical ports and of optical ports used and can be obtained by means of mathematical operations based on matrix algebra.

Computation of Network Cost

For an instance, $\Pi_{\mathcal{D}^E,\rho,\lambda} = \{(P_{1,\rho_1}, L_{1,\rho_1,\lambda_{1,\rho_1}}), (P_{2,\rho_2}, L_{2,\rho_2,\lambda_{2,\rho_2}}), \dots, (P_{M^E,\rho_{M^E}}, L_{M^E,\rho_{M^E},\lambda_{M^E,\rho_{M^E}}})\}$ of the network design problem under SED traffic requests, we define the following matrices:

- **Request Matrix:** The request matrix, denoted by $\underline{\Delta}^E = \{\delta_{i,t}^E\}$, represents the SED traffic requests over time. $\underline{\Delta}^E$ is a $[M^E \times \mathbb{E}]$ matrix. An element $\delta_{i,t}^E$ of such a matrix is a binary value ($\delta_{i,t}^E \in \{0, 1\}$) specifying the presence ($\delta_{i,t}^E = 1$) or absence ($\delta_{i,t}^E = 0$) of the SED request $\delta_i^E(s_i, d_i, \alpha_i, \beta_i, n_i)$ at time instant ϵ_t .

$$\delta_{i,t}^E = \begin{cases} 1 & \text{if } \alpha_i \leq \epsilon_t < \beta_i, \\ 0 & \text{otherwise.} \end{cases} \tag{16.2}$$

- **Source/Destination Matrix:** The source matrix, denoted by $\underline{S}^E = \{s_{i,n}^E\}$, represents the source nodes of the SED requests. The destination matrix, denoted by $\underline{D}^E = \{d_{i,n}^E\}$, represents the destination nodes of the SED requests. \underline{S}^E and \underline{D}^E are $[M^E \times N]$ matrices. An element $s_{i,n}^E$ ($d_{i,n}^E$) of such a matrix is a binary value specifying if the SED request $\delta_i^E(s_i, d_i, \alpha_i, \beta_i, n_i)$ has node v_n as source (destination) node or not.

$$s_{i,n}^E = \begin{cases} 1 & \text{if } s_i = v_n, \\ 0 & \text{otherwise.} \end{cases} \quad d_{i,n}^E = \begin{cases} 1 & \text{if } d_i = v_n, \\ 0 & \text{otherwise.} \end{cases} \quad (16.3)$$

As shown in Section 16.3.1, when grooming two SED requests $\delta_{i_1}^E$ and $\delta_{i_2}^E$, the set \mathcal{D}^E of all the SEDs must be modified by adding the aggregated request and the set of marginal demands, and by removing the original requests $\delta_{i_1}^E$ and $\delta_{i_2}^E$. Consequently, the number of SED requests as well as their characteristics are modified. Let $\tilde{\mathcal{D}}^E = \{\tilde{\delta}_1^E, \tilde{\delta}_2^E, \dots, \tilde{\delta}_{M^E}^E\}$ be the set of the \tilde{M}^E groomed SED requests obtained at the end of the grooming optimization process. Each groomed SED request $\tilde{\delta}_i^E \in \tilde{\mathcal{D}}^E$ is routed over a unique path \tilde{P}_i^E . As each element $\tilde{\delta}_i^E$ can be either a single low-speed connection δ_j^E or the concatenation of two or several low-speed connections $\tilde{\delta}_i^E = \uplus_j \delta_j^E$, its corresponding path is determined according to the grooming lightpath sets of its lower-speed connections ($\tilde{P}_i^E \in L_{j,\rho_j,\lambda_j,\rho_j}$).

Based on the new SED set $\tilde{\mathcal{D}}^E$, we define the following additional matrices:

- **Groomed Request Matrix:** The groomed request matrix, denoted by $\underline{\tilde{\Delta}}^E = \{\tilde{\delta}_{i,t}^E\}$, represents the groomed SED traffic requests over time. $\underline{\tilde{\Delta}}^E$ is a $[\tilde{M}^E \times \mathbb{E}]$ matrix. An element $\tilde{\delta}_{i,t}^E$ of such a matrix is an integer number due to the fact that the groomed SED demand $\tilde{\delta}_i^E(\tilde{s}_i, \tilde{d}_i, \tilde{\alpha}_i, \tilde{\beta}_i, \tilde{n}_i)$ can be the concatenation of several requests of smaller rates. Consequently, it could require more than one lightpath to carry its traffic load. $\tilde{\delta}_{i,t}^E$ is equal to the smallest integer greater than or equal to the traffic required by the groomed SED request $\tilde{\delta}_i^E$ and specifies the number of lightpaths used by this request at time instant ϵ_t .

$$\tilde{\delta}_{i,t}^E = \begin{cases} \lceil \tilde{n}_i \rceil & \text{if } \tilde{\alpha}_i \leq \epsilon_t < \tilde{\beta}_i, \\ 0 & \text{otherwise.} \end{cases} \quad (16.4)$$

- **Groomed Routing Matrix:** The groomed routing matrix, denoted by $\underline{\tilde{R}}^E = \{\tilde{r}_{i,f}^E\}$, represents the use of the physical links by the groomed SED requests. $\underline{\tilde{R}}^E$ is a $[\tilde{M}^E \times F]$ matrix. An element $\tilde{r}_{i,f}^E$ of such a matrix is a binary value ($\tilde{r}_{i,f}^E \in \{0, 1\}$) specifying if the physical path \tilde{P}_i^E assigned to the groomed SED request $\tilde{\delta}_i^E(\tilde{s}_i, \tilde{d}_i, \tilde{\alpha}_i, \tilde{\beta}_i, \tilde{n}_i)$ passes through the physical link e_f ($\tilde{r}_{i,f}^E = 1$) or not ($\tilde{r}_{i,f}^E = 0$).

$$\tilde{r}_{i,f}^E = \begin{cases} 1 & \text{if } e_f \in \tilde{P}_i^E, \\ 0 & \text{otherwise.} \end{cases} \quad (16.5)$$

- **Groomed Source/Destination Matrix:** The groomed source matrix, denoted by $\tilde{\underline{S}}^E = \{\tilde{s}_{i,n}^E\}$, represents the source nodes of the groomed SED requests. The groomed destination matrix, denoted by $\tilde{\underline{D}}^E = \{\tilde{d}_{i,n}^E\}$, represents the destination nodes of the groomed SED requests. $\tilde{\underline{S}}^E$ and $\tilde{\underline{D}}^E$ are $[\tilde{M}^E \times N]$ matrices. An element $\tilde{s}_{i,n}^E$ ($\tilde{d}_{i,n}^E$) of such a matrix is a binary value specifying if the groomed SED request $\tilde{\delta}_i^E(\tilde{s}_i, \tilde{d}_i, \tilde{\alpha}_i, \tilde{\beta}_i, \tilde{n}_i)$ has node v_n as source (destination) node or not.

$$\tilde{s}_{i,n}^E = \begin{cases} 1 & \text{if } \tilde{s}_i = v_n, \\ 0 & \text{otherwise.} \end{cases} \quad \tilde{d}_{i,n}^E = \begin{cases} 1 & \text{if } \tilde{d}_i = v_n, \\ 0 & \text{otherwise.} \end{cases} \quad (16.6)$$

By multiplying the transposed request matrix $\underline{\Delta}^{E^T}$ by the source matrix \underline{S}^E a new matrix $\underline{\Psi}^{E,Em} = \{\psi_{t,n}^{E,Em}\}$ is obtained. $\underline{\Psi}^{E,Em}$, of dimension $[\mathbb{E} \times N]$, represents the use of sub-wavelength ‘add’ ports at the nodes. An element $\psi_{t,n}^{E,Em}$ of such a matrix is an integer value representing the number of emitting e_1 electrical ports in use at node v_n at each instant ϵ_t . For a given node v_n , let $v_n^{E,Em}$ be the maximum value of the nodes’ add ports $\psi_{t,n}^{E,Em}$ evaluated over the whole observation period. $v_n^{E,Em}$ gives the number of emitting e_1 electrical ports installed at the node v_n . The number $\gamma_{e_1}^E$ of emitting e_1 electrical ports in the network is the sum of $v_n^{E,Em}$ over all the network nodes.

$$\underline{\Psi}^{E,Em} = \underline{\Delta}^{E^T} \times \underline{S}^E \quad (16.7a)$$

$$v_n^{E,Em} = \max_{1 \leq t \leq \mathbb{E}} \psi_{t,n}^{E,Em} \quad (16.7b)$$

$$\gamma_{e_1}^E = \sum_{n=1}^N v_n^{E,Em} \quad (16.7c)$$

Similarly, the number $\gamma_{r_1}^E$ of receiving r_1 electrical ports in the network is computed as follows:

$$\underline{\Psi}^{E,Re} = \underline{\Delta}^{E^T} \times \underline{D}^E \quad (16.8a)$$

$$v_n^{E,Re} = \max_{1 \leq t \leq \mathbb{E}} \psi_{t,n}^{E,Re} \quad (16.8b)$$

$$\gamma_{r_1}^E = \sum_{n=1}^N v_n^{E,Re} \quad (16.8c)$$

By multiplying the transposed groomed request matrix $\tilde{\Delta}^{E^T}$ by the groomed routing matrix \tilde{R}^E , a new matrix $\tilde{\Phi}^E = \{\tilde{\phi}_{t,f}^E\}$ is obtained. $\tilde{\Phi}^E$, of dimension $[\mathbb{E} \times F]$, represents the traffic load over time carried by the links. An element $\tilde{\phi}_{t,f}^E$ of such a matrix is an integer value representing the number of WDM optical channels carried by link e_f at each instant e_t . For a given link e_f , let $\tilde{\varphi}_f^E$ be the maximum value of this traffic $\tilde{\phi}_{t,f}^E$ evaluated over the whole observation period. $\tilde{\varphi}_f^E$ gives the number of WDM optical channels used on this link. As the o_1 optical ports go by pair, one port at each end of a WDM channel, $\tilde{\varphi}_f^E$ represents also the number of o_1 optical port pairs installed at both ends of the link e_f . The number $\gamma_{o_1}^E$ of o_1 optical ports in the network is twice the sum of $\tilde{\varphi}_f^E$ over all the network links. The maximum value of $\tilde{\varphi}_f^E$ ($\forall f \setminus 1 \leq f \leq F$) represents the congestion in the network, i.e., the number of active channels on the most loaded link.

$$\tilde{\Phi}^E = \tilde{\Delta}^{E^T} \times \tilde{R}^E \tag{16.9a}$$

$$\tilde{\varphi}_f^E = \max_{1 \leq t \leq \mathbb{E}} \tilde{\phi}_{t,f}^E \tag{16.9b}$$

$$\gamma_{o_1}^E = 2 \times \sum_{f=1}^F \tilde{\varphi}_f^E \tag{16.9c}$$

By multiplying the transposed groomed request matrix $\tilde{\Delta}^{E^T}$ by the groomed source matrix \tilde{S}^E a new matrix $\tilde{\Psi}^{E,Em} = \{\tilde{\psi}_{t,n}^{E,Em}\}$ is obtained. $\tilde{\Psi}^{E,Em}$, of dimension $[\mathbb{E} \times N]$, represents the use of switching ports from the EXC to the WXC at the nodes. An element $\tilde{\psi}_{t,n}^{E,Em}$ of such a matrix is an integer value representing the number of emitting e_3 electrical ports in use at node v_n at each instant e_t . For a given node v_n , let $\tilde{v}_n^{E,Em}$ be the maximum value of the nodes' switching ports $\tilde{\psi}_{t,n}^{E,Em}$ evaluated over the whole observation period. $\tilde{v}_n^{E,Em}$ gives the number of emitting e_3 electrical ports installed at the node v_n . The number $\gamma_{e_3}^E$ of emitting e_3 electrical ports in the network is the sum of $\tilde{v}_n^{E,Em}$ over all the network nodes.

$$\tilde{\Psi}^{E,Em} = \tilde{\Delta}^{E^T} \times \tilde{S}^E \tag{16.10a}$$

$$\tilde{v}_n^{E,Em} = \max_{1 \leq t \leq \mathbb{E}} \tilde{\psi}_{t,n}^{E,Em} \tag{16.10b}$$

$$\gamma_{e_3}^E = \sum_{n=1}^N \tilde{v}_n^{E,Em} \tag{16.10c}$$

Similarly, the number $\gamma_{r_3}^E$ of receiving r_3 electrical ports in the network is computed as follows:

$$\underline{\tilde{\Psi}}^{E,Re} = \underline{\tilde{\Delta}}^{E^T} \times \underline{\tilde{D}}^E \quad (16.11a)$$

$$\tilde{v}_n^{E,Re} = \max_{1 \leq t \leq \mathbb{E}} \tilde{\psi}_{t,n}^{E,Re} \quad (16.11b)$$

$$\gamma_{r_3}^E = \sum_{n=1}^N \tilde{v}_n^{E,Re} \quad (16.11c)$$

Consequently, the number $\gamma_{o_3}^E$ of o_3 optical ports in the network is equal to:

$$\gamma_{o_3}^E = \gamma_{e_3}^E + \gamma_{r_3}^E \quad (16.12)$$

As a set of SEDs does not have any SLD component to be added or dropped at a node, the number of e_2/r_2 electrical ports and the number of o_2 optical ports are null.

$$\gamma_{e_2}^E = \gamma_{r_2}^E = \gamma_{o_2}^E = 0 \quad (16.13)$$

Finally, the total number ϱ^E of optical ports and the total number ε^E of electrical ports are given by:

$$\begin{aligned} \varrho^E &= \gamma_{o_1}^E + \gamma_{o_2}^E + \gamma_{o_3}^E \\ &= \gamma_{o_1}^E + \gamma_{o_3}^E \end{aligned} \quad (16.14a)$$

$$\begin{aligned} \varepsilon^E &= \gamma_{e_1}^E + \gamma_{e_2}^E + \gamma_{e_3}^E + \gamma_{r_1}^E + \gamma_{r_2}^E + \gamma_{r_3}^E \\ &= \gamma_{e_1}^E + \gamma_{e_3}^E + \gamma_{r_1}^E + \gamma_{r_3}^E \end{aligned} \quad (16.14b)$$

κ being the ratio of the cost of an electrical port to the cost of an optical port, the cost ζ^E of routing the SED requests is then expressed as:

$$\zeta^E = \varrho^E + \kappa \times \varepsilon^E \quad (16.15)$$

16.4.2 Optical Grooming

$G = (V, E, w)$

is an arc-weighted symmetrical directed graph with vertex set $V = \{v_1, v_2, \dots, v_N\}$, arc set $E = \{e_1, e_2, \dots, e_F\}$, and arc weight function $w : E \rightarrow \mathbb{R}_+$. The graph represents a physical telecommunications network. The set V corresponds to the nodes of the network and the set of arcs E to the links interconnecting the nodes. Function w corresponds to the physical length or to the cost of the links (defined, for example, by the network operator).

$N = |V|$, $F = |E|$, K , L , B

are, respectively, the number of nodes and links in the network, the maximum number of possible alternate paths for each demand, the maximum number of possible layouts for each path and the size of a waveband-switching connection

(expressed in number of lightpaths). All the waveband-switching connections in the network have the same size.

$$P = (x_0, x_1, \dots, x_z)$$

describes a *path* in G that is composed of z links $(x_0, x_1), (x_1, x_2), \dots, (x_{z-1}, x_z)$ where the $(x_{i-1}, x_i) \in E$ are all distinct (i.e., the path is loop free).

$$\Delta = \{\delta_1, \delta_2, \dots, \delta_M\}$$

is a set of M Scheduled Lightpath Demands (SLDs), where:

$$\delta_i = (s_i, d_i, n_i, \alpha_i, \beta_i)$$

is a tuple representing SLD number i ; $s_i, d_i \in V$ are the source and destination nodes of the demand, n_i is the number of requested lightpaths, and α_i and β_i are the set-up and tear-down dates.

$$(G, \Delta)$$

is a pair representing an instance of the optical grooming problem.

$$P_{i,k} = (x_0^{i,k}, x_1^{i,k}, \dots, x_{z_i,k}^{i,k}), 1 \leq i \leq M, 1 \leq k \leq K$$

is the k -th alternate path for SLD δ_i from $x_0^{i,k} = s_i$ to $x_{z_i,k}^{i,k} = d_i$. For the purposes of this model, we compute the K physically⁴ shortest paths (if so many exist) for each demand using the algorithm defined in [14]. However, the paths might be defined according to any other criterion (i.e., the function w may map any other value than the length of the links).

$$C : \Pi_\Delta \rightarrow \mathbb{R}_+$$

is the function that computes the cost of an admissible solution $\pi_{\rho,\nu,\Delta}$ to the optical grooming problem (explained below). In order to formalize this function we define the following additional notations.

$$\theta = (\theta_{ij})$$

is a $\{0, 1\}^{M \times M}$ upper triangular matrix; θ_{ij} , $i \leq j$, indicates whether the SLDs δ_i and δ_j overlap in time ($\theta_{ij} = 1$) or not ($\theta_{ij} = 0$). By definition $\theta_{ii} = 1$, $1 \leq i \leq M$, and $\theta_{ij} = 0$ for $i > j$. This matrix expresses the temporal interdependence between the SLDs.

$$\tau = (\tau_{ij}) = \text{diag}(n_i)$$

is a diagonal matrix in which $\tau_{ii} = n_i$, $1 \leq i \leq M$, that is, τ_{ii} is the number of lightpaths required by the SLD δ_i .

$$\mathcal{I} = (\mathcal{I}_{ij})$$

is a $\{0, 1\}^{N \times F}$ matrix; \mathcal{I}_{ij} indicates whether vertex v_i is the termination vertex of arc e_j ($\mathcal{I}_{ij} = 1$) or not ($\mathcal{I}_{ij} = 0$).

⁴ The function w maps the length of each link.

$$\mathcal{O} = (\mathcal{O}_{ij})$$

is a $\{0, 1\}^{N \times F}$ matrix; \mathcal{O}_{ij} indicates whether vertex v_i is the source vertex of arc e_j ($\mathcal{O}_{ij} = 1$) or not ($\mathcal{O}_{ij} = 0$).

$$\mathcal{T} = (t_{ij})$$

is a $\{0, 1\}^{M \times N}$ matrix; t_{ij} indicates whether vertex v_j is the source node of SLD δ_i ($t_{ij} = 1$) or not ($t_{ij} = 0$).

$$\mathcal{U} = (u_{ij})$$

is a $\{0, 1\}^{M \times N}$ matrix; u_{ij} indicates whether vertex v_j is the destination node of SLD δ_i ($u_{ij} = 1$) or not ($u_{ij} = 0$).

$$\mathcal{D} = \theta \cdot \tau \cdot \mathcal{T} = (\mathcal{D}_{ij})_{1 \leq i \leq M, 1 \leq j \leq N},$$

$$\mathcal{D}_j^* = \max_{1 \leq i \leq M} \{\mathcal{D}_{ij}\}$$

\mathcal{D}^* is an N -dimensional vector; \mathcal{D}_j^* is the maximum number of simultaneously active lightpaths (i.e., time-overlapping lightpaths) originating at node v_i .

$$\mathcal{E} = \theta \cdot \tau \cdot \mathcal{U} = (\mathcal{E}_{ij})_{1 \leq i \leq M, 1 \leq j \leq N},$$

$$\mathcal{E}_j^* = \max_{1 \leq i \leq M} \{\mathcal{E}_{ij}\}$$

\mathcal{E}^* is an N -dimensional vector; \mathcal{E}_j^* is the maximum number of simultaneously active lightpaths (i.e., time-overlapping lightpaths) terminating at node v_i .

$$\psi : \mathbb{N} \rightarrow \mathbb{R}_+$$

is the function that determines the cost of a cross-connect (either a WXC or a BXC) as a function of its number of ports. In this work, we consider the function:

$$\psi(x) = a + bx^c \quad a, b \in \mathbb{R}_+, c \in [1, 2[. \quad (16.16)$$

The function captures various technology-specific factors that have an effect on the cost of a cross-connect. The parameter a represents the fixed cost of installing the switch (shelf, power and ventilation systems, etc.). The parameter b represents the cost of a port. Finally, the parameter c accounts for the impact on the cost of the increasing implementation complexity of switching matrices with a large number of ports. This parameter is limited to take values smaller than 2 because, for higher values, it is in general more economical to stack multiple small switches than building a large one when a significant number of ports is required.

$$\lambda = \{\lambda^n\}$$

is a partition of the set of arcs describing path P . We call λ a *layout* and an element λ^n of this partition a *subpath*. For example, a layout of path $(x_0, x_1, x_2, x_3, x_4)$ is $\lambda = \{(x_0, x_1, x_2), (x_2, x_3, x_4)\}$. This layout has subpaths $\lambda^1 = (x_0, x_1, x_2)$ and $\lambda^2 = (x_2, x_3, x_4)$. Another layout is $\{(x_0, x_1), (x_1, x_2, x_3, x_4)\}$. The elements of a subpath must be contiguous arcs of P .

$$\Lambda_{i,k} = \{\lambda_{i,k,j}\}, 1 \leq i \leq M, 1 \leq k \leq K, 1 \leq j \leq L$$

is the set of layouts available for path $P_{i,k}$; $\lambda_{i,k,j}$ is the j -th layout of the path. We assume that there are at most L different layouts defined for a path.

$$\pi_{\rho,\nu,\Delta} = ((P_{1,\rho_1}, \lambda_{1,\rho_1,\nu_1}), (P_{2,\rho_2}, \lambda_{2,\rho_2,\nu_2}), \dots, (P_{M,\rho_M}, \lambda_{M,\rho_M,\nu_M})),$$

$$\rho \in \{1, \dots, K\}^M, \nu \in \{1, \dots, L\}^M$$

is called an *admissible optical grooming solution* for Δ . ρ and ν are M -dimensional vectors whose elements can take values in $[1, K]$ and $[1, L]$, respectively. $\pi_{\rho,\nu,\Delta}$ is fully characterized by the pair (ρ, ν) . An admissible optical grooming solution $\pi_{\rho,\nu,\Delta}$ defines for each SLD, a path $P_{i,k}$ and a layout $\lambda_{i,k,j}$ of this path. A subpath λ^n can be part of several layouts defined in $\pi_{\rho,\nu,\Delta}$. We call the association of λ^n to the subset $\Delta' \subseteq \Delta$ of SLDs whose layouts in $\pi_{\rho,\nu,\Delta}$ share λ^n , a Routed Scheduled Band Group (RSBG). Thus, besides a path and a layout for each SLD, an admissible optical grooming solution $\pi_{\rho,\nu,\Delta}$ also defines a set of RSBGs.

$$\Pi_{\Delta} = \{\pi_{\rho,\nu,\Delta}, \rho \in \{1, \dots, K\}^M, \nu \in \{1, \dots, L\}^M\}$$

is the set of all admissible optical grooming solutions for Δ . Its cardinality is $|\Pi_{\Delta}| = (KL)^M$ (assuming that K paths and L layouts are available for each path; otherwise, $|\Pi_{\Delta}| < (KL)^M$). The set represents the solution space of an optical grooming problem instance (G, Δ) .

$$\mathcal{S}_{\rho,\nu,\Delta} = \bigcup_{i=1}^M \lambda_{i,\rho_i,\nu_i}$$

is the set of all subpaths used in the admissible optical grooming solution $\pi_{\rho,\nu,\Delta}$. The cardinality of the set is denoted $S = |\mathcal{S}_{\rho,\nu,\Delta}|$. For the sake of simplicity, we note \mathcal{S} instead of $\mathcal{S}_{\rho,\nu,\Delta}$.

$$\mathcal{A} = (a_{ij})$$

is a $\{0, 1\}^{M \times S}$ matrix; a_{ij} indicates whether SLD δ_i uses subpath \mathcal{S}_j ($a_{ij} = 1$) or not ($a_{ij} = 0$) in admissible optical grooming solution $\pi_{\rho,\nu,\Delta}$.

$$\mathcal{B} = (b_{ij})$$

is a $\{0, 1\}^{F \times S}$ arc-subpath incidence matrix; b_{ij} indicates whether arc e_i is part of subpath \mathcal{S}_j ($b_{ij} = 1$) or not ($b_{ij} = 0$).

$$\eta = \theta \cdot \tau \cdot \mathcal{A} \cdot \mathcal{B}^T = (\eta_{ij})$$

is a $\mathbb{N}^{M \times F}$ matrix; η_{ij} indicates the number of time-overlapping lightpaths on link e_j between SLD δ_i and SLDs $\delta_{i+1}, \delta_{i+2}, \dots, \delta_M$.

$$\eta^* = (\eta_j^*)_{1 \leq j \leq F}, \eta_j^* = \left\lceil \frac{1}{B} \max_{1 \leq i \leq M} \eta_{ij} \right\rceil$$

is an F -dimensional vector; η_j^* indicates the number of wavebands required on arc e_j .

$$\mathcal{IN} = \mathcal{I} \cdot \eta^*$$

is an N -dimensional vector; \mathcal{IN}_i indicates the number of input ports required in the BXC of node v_i to implement admissible optical grooming solution $\pi_{\rho, \nu, \Delta}$.

$$\mathcal{OUT} = \mathcal{O} \cdot \eta^*$$

is an N -dimensional vector; \mathcal{OUT}_i indicates the number of output ports required in the BXC of node v_i to implement admissible optical grooming solution $\pi_{\rho, \nu, \Delta}$.

$$\mathcal{G} = (g_{ij})$$

is a $\{0, 1\}^{N \times S}$ matrix; g_{ij} indicates whether vertex v_i is the source of subpath \mathcal{S}_j ($g_{ij} = 1$) or not ($g_{ij} = 0$).

$$\mathcal{H} = (h_{ij})$$

is a $\{0, 1\}^{N \times S}$ matrix; h_{ij} indicates whether vertex v_i is the termination of subpath \mathcal{S}_j ($h_{ij} = 1$) or not ($h_{ij} = 0$).

$$\mathcal{J} = \theta \cdot \tau \cdot \mathcal{A} \cdot \mathcal{G}^T = (\mathcal{J}_{ij})_{1 \leq i \leq M, 1 \leq j \leq N},$$

$$\mathcal{J}_j^* = \left\lceil \frac{1}{B} \max_{1 \leq i \leq M} \mathcal{J}_{ij} \right\rceil$$

\mathcal{J}^* is an N -dimensional vector; \mathcal{J}_j^* is the number of waveband connections added at the BXC of node v_j .

$$\mathcal{K} = \theta \cdot \tau \cdot \mathcal{A} \cdot \mathcal{H}^T = (\mathcal{K}_{ij})_{1 \leq i \leq M, 1 \leq j \leq N},$$

$$\mathcal{K}_j^* = \left\lceil \frac{1}{B} \max_{1 \leq i \leq M} \mathcal{K}_{ij} \right\rceil$$

\mathcal{K}^* is an N -dimensional vector; \mathcal{K}_j^* is the number of waveband connections dropped at the BXC of node v_j .

$$\mathcal{Q} \in \mathbb{R}_+^N,$$

$$\mathcal{Q}_i = \psi_b(\max(\mathcal{IN}_i, \mathcal{OUT}_i) + \max(\mathcal{J}_i^*, \mathcal{K}_i^*))$$

is an N -dimensional vector; \mathcal{Q}_i indicates the cost of the BXC required at node v_i to implement admissible optical grooming solution $\pi_{\rho, \nu, \Delta}$.

$$\mathcal{R} \in \mathbb{R}_+^N,$$

$$\mathcal{R}_i = \psi_w(\max(\mathcal{D}_i^*, \mathcal{E}_i^*) + B \max(\mathcal{J}_i^*, \mathcal{K}_i^*))$$

is an N -dimensional vector; \mathcal{R}_i indicates the cost of the WXC required at node v_i to implement admissible optical grooming solution $\pi_{\rho, \nu, \Delta}$. Note that the number of input/output ports, $B \max(\mathcal{J}_i^*, \mathcal{K}_i^*)$, is a multiple of the waveband size B . The parameters defining the function ψ_w may be different from those defining the function ψ_b .

The cost function $C : \Pi_{\Delta} \rightarrow \mathbb{R}_+$ is defined as:

$$C(\pi_{\rho, \nu, \Delta}) = \sum_{i=1}^N (Q_i + R_i). \tag{16.17}$$

16.5 Illustrative Examples

16.5.1 Electrical Grooming

In this subsection, we present a small example in order to highlight the reduction of network cost provided by the electrical grooming functionality. We considered the three 3 SEDs of Table 16.1. Their time diagram is shown in Fig. 16.8. To emphasize the benefit of grooming we use a fixed routing solution in which each SED is routed along its shortest path. Emitting e_1 electrical ports and receiving r_1 electrical ports are only needed to add and drop the traffic requests. Consequently, the number of such ports is independent of the routing and grooming solution. When the network does not have grooming capabilities (Fig. 16.9(a)), the SEDs δ_1 and δ_2 are routed over distinct lightpaths. Thus, we need two WDM optical channels on each of the 3–4 and the 4–7 fiber links. As a result, this routing and grooming solution requires a total of 9 WDM optical channels. Let us recall that the number of required o_1 optical ports is twice the number of WDM optical channels used. On the other hand, when the network has grooming capabilities (Fig. 16.9(b)), the SEDs δ_1 and δ_2 can be groomed together since both requests are active from 11:00 to 13:00 and they share a common path between nodes 3 and 7. Therefore, a grooming lightpath is created between these nodes using only one WDM optical channel on each of the 3–4 and the 4–7 fiber links. This grooming

Table 16.1. Characteristics of 3 SEDs

δ_i	s_i	d_i	α_i	γ_i	n_i
δ_1	2	8	8:00	14:40	$0.25 \times C_{\omega}$
δ_2	3	7	11:00	13:00	$0.25 \times C_{\omega}$
δ_3	2	6	17:00	19:30	$0.25 \times C_{\omega}$

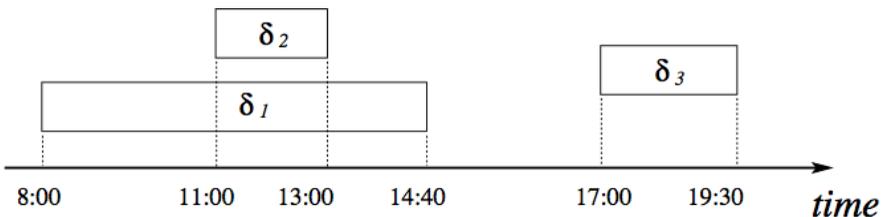


Fig. 16.8. Associated time diagram of the 3 SEDs

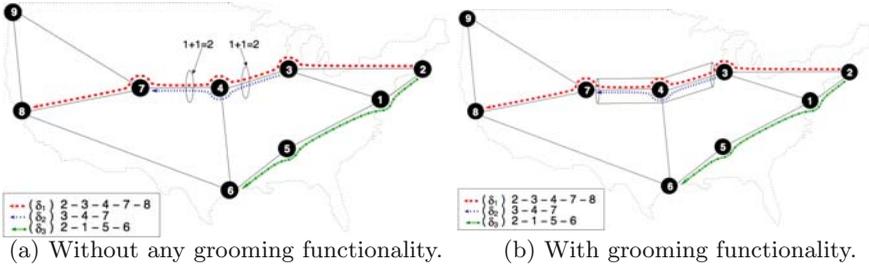


Fig. 16.9. Provisioning in a network of the 3 SEDs of Table 16.1

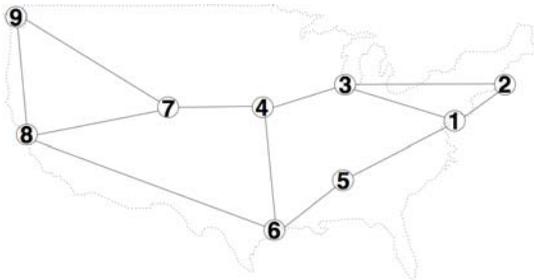
lightpath is used to carry simultaneously the two SEDs δ_1 and δ_2 between nodes 3 and 7. In our example, the SED δ_1 still needs to be routed using two additional lightpaths; the first one between nodes 2 and 3, and the second one between nodes 7 and 8. As a result, only 7 WDM optical channels are needed for this new routing and grooming solution. However, it should be noted that, in order to aggregate SEDs δ_1 and δ_2 at node 3, one additional emitting o_3 optical port and one additional receiving r_3 electrical port are needed on this node. Similarly, to de-aggregate these two SEDs at node 7, one additional emitting e_3 electrical port and one additional receiving o_3 optical port are needed. Depending on the relative costs of the different o_1 , o_3 , e_3 , and r_3 ports, the new routing and grooming solution can be more or less economical than the first solution.

16.5.2 Optical Grooming

In this subsection, we present an example to illustrate how the mathematical model of Section 16.4.2 is used to represent an instance of the optical grooming problem and to quantify the cost of a solution to this problem.

We study the problem instance of Fig. 16.10 which defines a set Δ of $M = 2$ SLDs and a graph G representing a physical network. Additionally, we choose the paths and layouts defined in Table 16.2 ($K = 2$ and $L = 3$) and define a band size of $B = 4$. For the sake of simplicity, we choose functions $\psi_w(x) = x$ and $\psi_b(x) = x$ to represent the cost of switches, i.e., we choose $a = 0$, $b = 1$ and $c = 1$ in Equation (16.16). Figure 16.11 illustrates the admissible optical grooming solution $\pi_{\rho,\nu,\Delta}$ with vectors $\rho = (1, 1)$ and $\nu = (3, 3)$ for the instance of Fig. 16.10. Path $P_{1,1}$ and layout $\lambda_{1,1,3}$ are selected for SLD δ_1 and path $P_{2,1}$ and layout $\lambda_{2,1,3}$ are selected for SLD δ_2 . The figure also shows the set of Routed Scheduled Band Groups (RSBGs) defined by the solution.

The following matrices are used to compute the cost of the chosen admissible optical grooming solution:



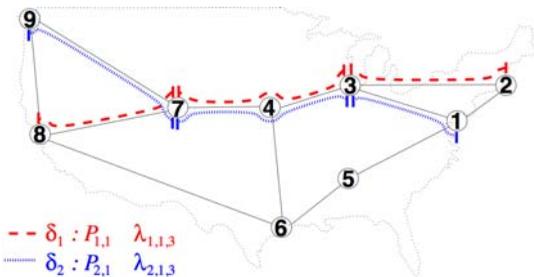
SLD	s	d	n	α	β
δ_1	2	8	8	08:00	14:00
δ_2	1	9	10	11:00	13:00

(a) Graph G representing a physical network. (b) Set Δ of $M = 2$ SLDs.

Fig. 16.10. An instance (G, Δ) of the optical grooming problem

Table 16.2. Chosen paths and layouts for the instance of Fig. 16.10

SLD	Path ($P_{i,k}$)	Layout ($\lambda_{i,k,j}$)
δ_1	$P_{1,1} = (2, 3, 4, 7, 8)$	$\lambda_{1,1,1} = \{(2, 3, 4, 7, 8)\}$
		$\lambda_{1,1,2} = \{(2, 3, 4) (4, 7, 8)\}$
		$\lambda_{1,1,3} = \{(2, 3) (3, 4, 7) (7, 8)\}$
	$P_{1,2} = (2, 1, 5, 6, 8)$	$\lambda_{1,2,1} = \{(2, 1, 5, 6, 8)\}$
		$\lambda_{1,2,2} = \{(2, 1, 5) (5, 6, 8)\}$
		$\lambda_{1,2,3} = \{(2, 1) (1, 5, 6) (6, 8)\}$
δ_2	$P_{2,1} = (1, 3, 4, 7, 9)$	$\lambda_{2,1,1} = \{(1, 3, 4, 7, 9)\}$
		$\lambda_{2,1,2} = \{(1, 3, 4) (4, 7, 9)\}$
		$\lambda_{2,1,3} = \{(1, 3) (3, 4, 7) (7, 9)\}$
	$P_{2,2} = (1, 5, 6, 8, 9)$	$\lambda_{2,2,1} = \{(1, 5, 6, 8, 9)\}$
		$\lambda_{2,2,2} = \{(1, 5, 6) (6, 8, 9)\}$
		$\lambda_{2,2,3} = \{(1, 5) (5, 6, 8) (8, 9)\}$



Subpath (λ^n)	$\Delta' \subseteq \Delta$
(1, 3)	$\{\delta_2\}$
(2, 3)	$\{\delta_1\}$
(3, 4, 7)	$\{\delta_1, \delta_2\}$
(7, 8)	$\{\delta_1\}$
(7, 9)	$\{\delta_2\}$

-- $\delta_1 : P_{1,1} \lambda_{1,1,3}$
 --- $\delta_2 : P_{2,1} \lambda_{2,1,3}$

(a) Paths and layouts.

(b) Routed Scheduled Band Groups (RSBGs).

Fig. 16.11. An admissible optical grooming solution $\pi_{\rho,\nu,\Delta}$ to the instance of Fig. 16.10

$$\theta = \begin{matrix} & \delta_1 & \delta_2 \\ \delta_1 & \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \end{matrix} \quad \gamma = \begin{matrix} & \delta_1 & \delta_2 \\ \delta_1 & \begin{pmatrix} 8 & 0 \\ 0 & 10 \end{pmatrix} \end{matrix}$$

$$\mathcal{S}_{\rho, \nu, \Delta} = \{(1, 3), (2, 3), (3, 4, 7), (7, 8), (7, 9)\}$$

$$\mathcal{A} = \begin{matrix} & s_1 & s_2 & s_3 & s_4 & s_5 \\ \delta_1 & \begin{pmatrix} 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \end{pmatrix} \end{matrix} \quad \mathcal{B} = \begin{matrix} & s_1 & s_2 & s_3 & s_4 & s_5 \\ e_{13} & \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix}$$

$$\eta = \theta \cdot \gamma \cdot \mathcal{A} \cdot \mathcal{B}^T = \begin{pmatrix} 10 & 8 & 18 & 18 & 8 & 10 \\ 10 & 0 & 10 & 10 & 0 & 10 \end{pmatrix} \quad \eta^* = \begin{pmatrix} 3 \\ 2 \\ 5 \\ 5 \\ 2 \\ 3 \end{pmatrix}$$

$$\mathcal{I} = \begin{matrix} & e_{13} & e_{23} & e_{34} & e_{47} & e_{78} & e_{79} \\ v_1 & \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix} \quad \mathcal{O} = \begin{matrix} & e_{13} & e_{23} & e_{34} & e_{47} & e_{78} & e_{79} \\ \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \end{matrix}$$

$$\mathcal{IN} = \mathcal{I} \cdot \eta^* = (0, 0, 5, 5, 0, 0, 5, 2, 3)$$

$$\mathcal{OUT} = \mathcal{O} \cdot \eta^* = (3, 2, 5, 5, 0, 0, 5, 0, 0)$$

$$\mathcal{G} = \begin{matrix} & s_1 & s_2 & s_3 & s_4 & s_5 \\ v_1 & \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \end{matrix} \quad \mathcal{H} = \begin{matrix} & s_1 & s_2 & s_3 & s_4 & s_5 \\ v_1 & \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix}$$

$$\begin{aligned} \mathcal{T} &= \begin{matrix} & v_1 & v_2 & v_3 & \dots & v_9 \\ \delta_1 & \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 1 & 0 & 0 & \dots & 0 \end{pmatrix} & \delta_2 & \begin{pmatrix} v_1 & \dots & v_7 & v_8 & v_9 \\ 0 & \dots & 0 & 1 & 0 \\ 0 & \dots & 0 & 0 & 1 \end{pmatrix} \end{matrix} \\ \mathcal{D} &= \theta \cdot \gamma \cdot \mathcal{T} = \begin{pmatrix} 10 & 8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 10 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad \mathcal{D}^* = (10 \ 8 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0) \\ \mathcal{E} &= \theta \cdot \gamma \cdot \mathcal{U} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 8 & 10 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 10 \end{pmatrix} \quad \mathcal{E}^* = (0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 8 \ 10) \\ \mathcal{J} &= \theta \cdot \gamma \cdot \mathcal{A} \cdot \mathcal{G}^T = \begin{pmatrix} 10 & 8 & 18 & 0 & 0 & 0 & 18 & 0 & 0 \\ 10 & 0 & 10 & 0 & 0 & 0 & 10 & 0 & 0 \end{pmatrix} \quad \mathcal{J}^* = (3 \ 2 \ 5 \ 0 \ 0 \ 0 \ 5 \ 0 \ 0) \\ \mathcal{K} &= \theta \cdot \gamma \cdot \mathcal{A} \cdot \mathcal{H}^T = \begin{pmatrix} 0 & 0 & 18 & 0 & 0 & 0 & 18 & 8 & 10 \\ 0 & 0 & 10 & 0 & 0 & 0 & 10 & 0 & 10 \end{pmatrix} \quad \mathcal{K}^* = (0 \ 0 \ 5 \ 0 \ 0 \ 0 \ 5 \ 2 \ 3) \\ \mathcal{Q} &= (6 \ 4 \ 10 \ 5 \ 0 \ 0 \ 10 \ 4 \ 6) \quad \mathcal{R} = (22 \ 16 \ 20 \ 0 \ 0 \ 0 \ 20 \ 16 \ 22) \end{aligned}$$

The cost of the admissible optical grooming solution $\pi_{\rho,\nu,\Delta}$ according to (16.17) is:

$$C(\pi_{\rho,\nu,\Delta}) = 28 + 20 + 30 + 5 + 0 + 0 + 30 + 20 + 28 = 161. \quad (16.18)$$

16.6 Conclusion

This chapter presented mathematical formulations to quantify the cost of provisioning scheduled connections in multi-layer optical networks. The specific cases of electrical grooming in EXC/WXC networks and optical grooming in WXC/BXC networks were considered. The formulations may serve as a basis to assess the costs and benefits of introducing an additional network layer in an existing network, or to determine the impact on the network-wide cost of a change in the technologies used to implement a cross-connect. The use of scheduled connections make it possible to take into account dynamic changes on the traffic load in the cost/benefit analysis of multi-layer networks.

The formulations could be extended to compare the impact on network cost of alternative resiliency approaches such as link or path protection, dedicated or shared protection, etc. Another extension could be the consideration of the uncertainty about the tear-down date of SEDs or SLDs, since connections terminating later than their expected tear-down date might be seizing network resources allocated to demands that were originally time-disjoint with respect to the seizing connection. It is reasonable to expect that, the longer a connection lasts in the network, the higher the uncertainty about its actual tear-down date.

References

1. J. Kuri, N. Puech, M. Gagnaire, E. Dotaro and R. Douville, "Routing and Wavelength Assignment of Scheduled Lightpath Demands," *IEEE Journal on Selected Areas in Communications*, Vol. 21, N. 8, pp. 1231–1240, October 2003.
2. J. Kuri, "Optimization Problems in WDM Optical Transport Networks with Scheduled Lightpath Demands," *Ph.D. Thesis*, Ecole Nationale Supérieure des Télécommunications. Paris, France. September 2003.
3. N. Skorin-Kapov, "Heuristic algorithms for the routing and wavelength assignment of scheduled lightpath demands in optical networks," *IEEE Journal on Selected Areas in Communications*, Vol. 24, NO. 8, August 2006.
4. H. G. Ahn, T.-J. Lee, M. Y. Chung and H. Choo, "RWA on Scheduled Lightpath Demands in WDM Optical Transport Networks with Time Disjoint Paths," *Proceedings of the International Conference on Information Networking (ICOIN 2005)*, Jeju, Korea, February 2005.
5. C. V. Saradhi and M. Gurusami, "Graph Theoretic Approaches for Routing and Wavelength Assignment of Scheduled Lightpath Demands in WDM Optical Networks," *Proceedings of the 2nd International Conference on Broadband Networks (Broadnets 2005)*, Boston, MA, October 2005.
6. J. Kuri, N. Puech, M. Gagnaire and E. Dotaro, "Routing and Wavelength Assignment of Scheduled Lightpath Demands in a WDM Optical Transport Network," *Proceedings of the First International Conference in Optical Communications and Networks (ICOON 2002)*, Singapur, November 2002.
7. T. Li and B. Wang, "On optimal survivability design in WDM optical networks under a scheduled traffic model," *Proceedings of the 5th International Workshop on Design of Reliable Communication Networks (DRCN 2005)*, Ischia, Italy, October 2005.
8. T. Li and B. Wang, "On Optimal Survivability Design under a Scheduled Traffic Model in Wavelength-Routed Optical Mesh Networks," *Proceedings of the 4th Annual Communication Networks and Services research Conference (CNSR 2006)*, Moncton, Canada, May 2006.
9. C. V. Saradhi, L. K. Wei and M. Gurusami, "Provisioning Fault-Tolerant Scheduled Lightpath Demands in WDM Mesh Networks," *Proceedings of the First International Conference on Broadband Networks (BROADNETS 2004)*, San Jose, CA, October 2004.
10. E. A. Doumith and M. Gagnaire, "Traffic Routing in Multi-Layer Optical Network Considering Rerouting and Grooming Strategies," *Proceedings of the IEEE Global Communications Conference (Globecom)*, San Francisco, CA, December 2006.
11. E. A. Doumith, M. Gagnaire, O. Audouin, and R. Douville, "Network Nodes Dimensioning Assuming Electrical Traffic Grooming in an Hybrid OXC/EXC WDM Network," *Proceedings of the 2nd IEEE/Create-Net International Conference on Broadband Networks (BroadNets'05)*, Boston, MA, 2005.
12. E. A. Doumith and M. Gagnaire, "Traffic Grooming in Multi-Layer WDM Networks: Meta-heuristics versus Sequential Algorithms," *Proceedings of the IEEE Eunice Open European Summer School*, Stuttgart, Germany, September 2006.

13. E. A. Doumith, M. Gagnaire, O. Audouin and R. Douville, "From Network Planning to Traffic Engineering for Optical VPN and Multi-Granular Random Demands," *Proceedings of the 2nd IEEE/Create-Net International Conference on Broadband Networks (BroadNets'05)*, Phoenix, AZ, April 2006.
14. D. Eppstein, "Finding the k Shortest Paths," *IEEE Symposium on Foundations of Computer Science*, pp. 154–165, 1994.

All-Optical Traffic Grooming

Volkan Kaman, Henrik N. Poulsen, Roger J. Helkey, John E. Bowers

17.1 Introduction

Next-generation optical communication networks are expected to exceed aggregate capacities of hundreds of terabits per second to support the recent growth in commercial broadband access fueled by applications such as Video-on-Demand (VoD) and Voice-over-IP (VoIP), and to meet the high-performance demands of grid-computing networks for research consortiums and military applications [15, 26, 31, 32]. The continuous increase in need for higher bandwidth has also necessitated in parallel a continuous reduction in cost per bit, which has paved the way for the all-optical revolution. As higher bandwidth with smaller footprint and lower power consumption scales better with lower cost optics, the past 30 years have seen electronics get gradually pushed out of the core communication network to the lower bandwidth edge user interfaces. The advantage of transitioning to all-optical communication is most evident first with the introduction of the erbium-doped fiber amplifier (EDFA) in the early 1990s. The EDFA not only increased the all-optical reach of bandwidth independent signals by eliminating costly optical-electrical-optical (OEO) repeaters, but it also revolutionized point-to-point communication system capacity by allowing widespread implementation of dense wavelength division multiplexing (DWDM). The second confirmation of the benefit of all-optical communication is the introduction of optical add-drop multiplexers (OADM) at metro and backbone network nodes in the early 2000s. Similarly to the EDFA, the OADM allowed the removal of costly in-line OEO transponders by providing all-optical bypass of express traffic. Furthermore, in realizing the concept of the all-optical core ring network, the OADM provides significant optics savings by driving the grooming electronics switch to the edge of the network for sub-wavelength processing of the local add-drop traffic.

While the EDFA and OADM have allowed for the realization of today's all-optical point-to-point systems and core ring networks, the current shift to higher data rates of 40 Gb/s SONET, 100 Gb Ethernet and beyond opens up the possibility for further cost-saving all-optical innovations in high-capacity

edge and core networks. This chapter reviews the next-generation all-optical traffic grooming technologies for both the core and edge networks. Section 17.2 reviews the state-of-the-art smart optical switching technologies for all-optical wavelength grooming in the core network as these networks evolve to interconnected rings and mesh architectures. Section 17.3 reviews all-optical circuit grooming, or optical time division multiplexing (OTDM), for aggregating low-speed sub-wavelength circuits into wavelengths between the core and edge networks. Finally, Section 17.4 reviews all-optical edge grooming techniques such as optical packet switching (OPS) and optical burst switching (OBS) for reducing costly electronics in edge grooming switches and IP routers, and Section 17.5 concludes the chapter.

17.2 All-Optical Wavelength Grooming

The aggregate bandwidth capacity increase in next-generation metro and long haul core optical networks requires a considerable increase in DWDM fiber link capacity, which is achieved by increasing both the number of wavelengths and the data rate per wavelength using spectrally efficient modulation formats as recently demonstrated with a 25.6 Tb/s transmission experiment [6]. These high-capacity DWDM links are interconnected into dynamically reconfigurable all-optical mesh networks with efficient wavelength provisioning and optimized use of network elements for both capital and operational cost savings. This agile photonic network is enabled by the next-generation of remotely reconfigurable OADM (ROADM) with express traffic switching and local add-drop wavelength grooming capabilities. The ROADM should support multiple degree nodes, which is defined as the number of DWDM fiber network ports terminating at the node. It should also allow each add-drop port to have fully flexible access to the network to ensure efficient network utility of each access transponder. The latter is of high significance in a multi-degree ROADM design as transponders and regenerators dominate the overall core network cost and should be minimized to the lowest possible required by the network [5, 31]. The key enablers for all-optical wavelength grooming of core network traffic are the dramatic savings in (i) cost per express wavelength switching in the optical domain by removing in-line OEO transponders and electronic switches, and (ii) cost per access wavelength switching by removing redundant edge OEO transponders through flexible optical add-drop.

Traditional degree-two nodes in ring networks have relied upon static OADMs that can transit and add-drop only fixed wavelengths without reconfigurability. These OADMs are either based on a three-port WDM filter approach (Fig. 17.1a) or a pair of concatenated optical demultiplexers (DMUX) and multiplexers (MUX) for splitting and recombining all DWDM wavelengths (Fig. 17.1b). Next-generation degree-two ROADMs are based on wavelength blockers (WB) using liquid crystals (LC), or one-dimensional micro-electro-mechanical-system (1-D MEMS) switches, or integrated planar

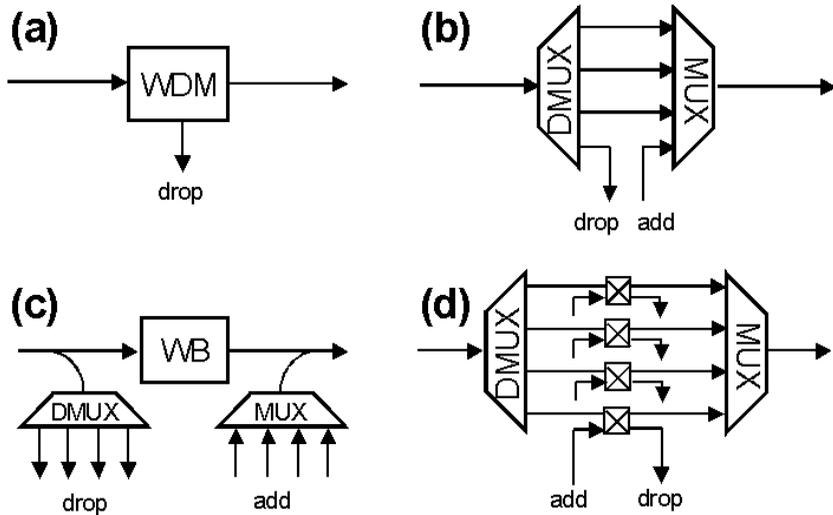


Fig. 17.1. Schematics of degree-two static OADM's based on (a) a WDM filter, (b) concatenated DMUX and MUX; and ROADM's based on (c) WB's, and (d) concatenated DMUX and MUX with 2×2 switches for add-drop access

lightwave circuits (PLC), where all techniques process and power balance express traffic at a per wavelength basis [18, 23]. Add-drop access for both 1-D MEMS and LC-based ROADMs is achieved by tapping the aggregate DWDM signal before and after the WB and separating the channels using fixed wavelength D/MUXs (Fig. 17.1c). The PLC-based ROADM has the advantage of utilizing the already integrated express D/MUXs for fixed wavelength add-drop access by further integrating per channel 2×2 switches onto the PLC device (Fig. 17.1d). While degree-two ROADMs provide reconfigurable all-optical wavelength grooming, fully flexible add-drop access requires the use of additional tunable filters and optical switches, the cost of which can be prohibitive for high add-drop ratios.

As next-generation networks evolve into interconnected rings and mesh architectures, multi-degree ROADM nodes with flexible add-drop capability gain high significance. The add-drop flexibility criteria that eliminate additionally redundant and expensive electronics with cheap optics are summarized as follows [10]:

1. *directionless access* with any add-drop port to any DWDM network port connectivity,
2. *colorless access* with any wavelength-transparent add-drop port to any DWDM network port connectivity, and
3. *contentionless access* between all same-wavelength add-drop ports.

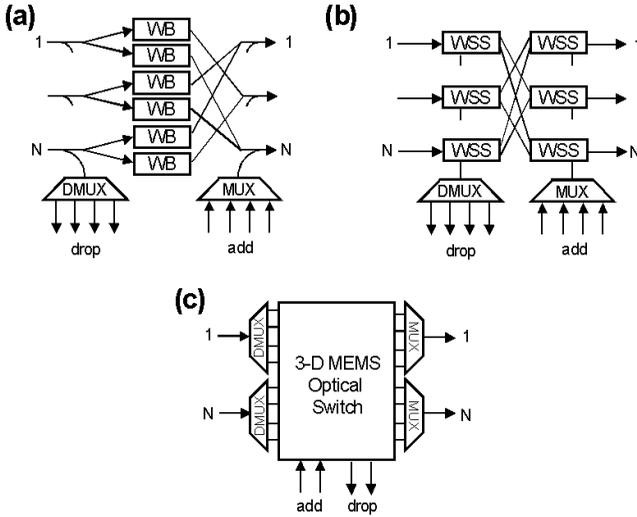


Fig. 17.2. Schematics of multi-degree ROADM architectures (a) broadcast-and-select using WBs, (b) optical cross-connect using WSSs, (c) WSXC using 3-D MEMS optical switch

Multi-degree ROADMs can be realized using three distinct technologies as summarized in Fig. 17.2: (a) a broadcast-and-select architecture using WBs, (b) an optical cross-connect (or broadcast-and-select) architecture using wavelength selective switches (WSS), and (c) a wavelength-selective cross-connect (WSXC) architecture using a large-scale 3-D MEMS optical switch.

The broadcast-and-select architecture shown in Fig. 17.2a is based on the same 1-D MEMS or LC WB technologies as discussed for degree-two ROADMs [30]. While passive optical splitters are used to broadcast and combine each incoming and outgoing DWDM network port, the WBs select the desired wavelengths to be transmitted. An inherent disadvantage of this architecture is that as the degree of the node N increases, the required number of WBs increases quadratically to $N(N - 1)$ while the ROADM express through loss increases significantly as well. This multi-degree ROADM also physically separates the express wavelength switching from the add-drop switching by tapping the DWDM signals before and after the core express switch. While the degree based modularity of this ROADM is attractive from a network restoration perspective, this also severely complicates the ability to provide flexible add-drop access with the aforementioned flexibility criteria.

A more attractive alternative to WB's is to use 1-D MEMS or LC based $1 : K$ WSS devices (typically with $K < 10$), which distribute an incoming DWDM signal into any of the desired K outputs [2, 20]. A multi-degree ROADM node can be achieved by cascading two WSS devices per network port and inter-connecting $N - 1$ WSS ports from one network port to the others

as shown in Fig. 17.2b [1]. While any of the remaining $K - N + 1$ WSS ports can be used for colorless local add-drop access of a single transponder, the limited number of available WSS ports requires optical aggregation (using D/MUX, splitters, and EDFAs) of the transponders before accessing the WSS. Furthermore, directionless and wavelength contentionless add-drop access requires further optical processing at the edge, which becomes undesirable for high-degree nodes with high add-drop ratios [13]. While the preceding described the optical cross-connect architecture using $2N$ WSS devices with a fixed through loss, typical ROADM implementations use the broadcast-and-select architecture by cascading an input splitter with a single WSS device per network port [1]. This architecture has the advantage that the number of required WSS devices per node is reduced to N ; however, the challenges with flexible add-drop access remain with an additional express insertion loss drawback.

A multi-degree ROADM can also be implemented with a WSXC architecture using a large-scale 3-D MEMS optical switch combined with a set of D/MUXs (Fig. 17.2c) [14]. While non-blocking 3-D MEMS switches with low optical insertion losses are commercially available with 320 ports, these matrix switches have also been shown to scale beyond 1,000 ports [9, 17]. Wavelength switching between DWDM network ports is achieved by directing each wavelength separately to a MEMS port through the use of modular or integrated D/MUX at the DWDM network ports. Furthermore, fully agile add-drop access is realized due to the colorless and non-blocking nature of each MEMS port. A transponder or regenerator can connect to the entire core network at any wavelength using a single MEMS port. Since the add-drop ports use the already existing express filtering devices, no additional elements are required for the add-drops, reducing cost compared to the preceding WSS architectures.

The three multi-degree ROADM technologies are compared in Fig. 17.3 in terms of WSS functionality and evolution. The first generation WSS is the WB with a single input and output. The limited number of ports on this 1:1 WSS not only results in a quadratic need of these devices in implementing a multi-degree ROADM, but also the absence of local access ports requires additional components with limited add-drop flexibility. The next-generation device is the 1: K WSS, which has a single input port and several output ports. The number of devices required to implement a multi-degree ROADM scales linearly with the number of network ports. While a limited number of the K ports can be used for colorless local access, full add-drop flexibility and high add-drop ratios requires additional components and complicated access architectures. Finally, the integrated WSXC realizes the multi-degree ROADM through a $N:N$ WSS device with additional ports available for fully flexible add and drop access at wavelength granularity with simple operational management. While the latter approach provides the most flexible multi-degree ROADM due to the non-blocking switching nature of each wavelength port, the former two architectures provide a modular growth possibility with less add-drop flexibility. While modular growth of 3-D MEMS switches has been proposed, it can be concluded that network port modularity and

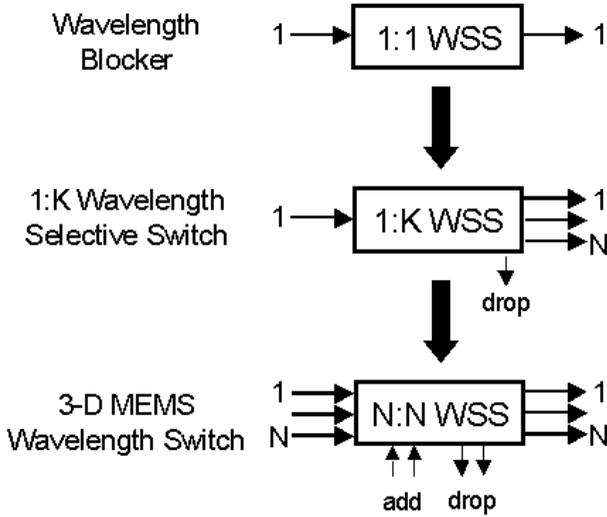


Fig. 17.3. Comparison of multi-degree ROADMs technologies in terms of WSS functionality and evolution

local add-drop flexibility of a multi-degree ROADM are competing design criteria [13].

17.3 All-Optical Circuit Grooming

While DWDM is widely utilized to upgrade fiber bandwidth capacity, the cost, provisioning, and management of a huge number of wavelength channels can be highly problematic when considering dynamic mesh networks as opposed to point-to-point systems. Increasing the bit rate of each wavelength channel using traditional serial TDM circuit grooming while keeping the number of DWDM wavelengths to a cost effective and reasonably manageable number is therefore required to satisfy the bandwidth requirements of next-generation high-capacity networks. While TDM based on electronic grooming is currently being deployed at 40 Gb/s in commercial networks and demonstrated at 100 Gb/s in lab trials [19, 29], the speed and cost limitations of electronics will most likely require all-optical circuit grooming or OTDM techniques for achieving higher data rates of 160 Gb/s and beyond per wavelength [33]. As sub-wavelength multiplexing and demultiplexing are performed in the optical domain, OTDM systems have the advantage of needing opto-electronic component bandwidths only at the lower speed tributary rate. For example, a 160 Gb/s data rate can be achieved with commercially available low-cost 10 GHz devices [24]. Optical grooming or aggregation capability between the edge and core networks also allows for efficient utilization of resources for dynamic

sub-wavelength traffic. Furthermore, as lower speed circuits within a given wavelength may have different destinations, the ability to drop the desired circuits and insert new circuits in the available slots, or add-drop multiplexing, becomes a desirable functionality [27]. While traditional TDM requires electronic processing of all sub-wavelength traffic within a given wavelength, optical add-drop multiplexing using OTDM techniques can provide complete optical transparency to express sub-wavelength traffic from source to destination.

The basic principle of OTDM is illustrated in a point-to-point system in Fig. 17.4a. A typical OTDM transmitter consists of an optical pulse source that emits short pulses at the desired wavelength with a repetition rate of T , which is determined by the bit rate B at which electronics can follow. A passive $1:M$ splitter broadcasts these pulses into separate paths, where each pulse stream is independently data modulated and time delayed. All M pulse streams are then bit interleaved into a single serial data stream with an aggregate data rate of $B \times M$. The number of serial channels that can be packed into a single wavelength depends on the width of the optical pulses. On the OTDM receiver side, a $1:M$ splitter broadcasts the aggregate data stream to each of the M tributary detectors. A phase-locked optical gating device (DMUX) is utilized prior to the detector for extracting the desired sub-wavelength channel. The concept of add-drop multiplexing using the aforementioned OTDM techniques is shown in Fig. 17.4b. An optical gate extracts the desired channel to be dropped at the node while the rest of the channels are transmitted transparently. A new channel is then inserted in the available time slot.

While various OTDM technologies have been demonstrated in lab trials based on all-optical fiber techniques for ultra-high-speed single wavelength data rates [25, 33], the high number of devices required in OTDM sub-systems

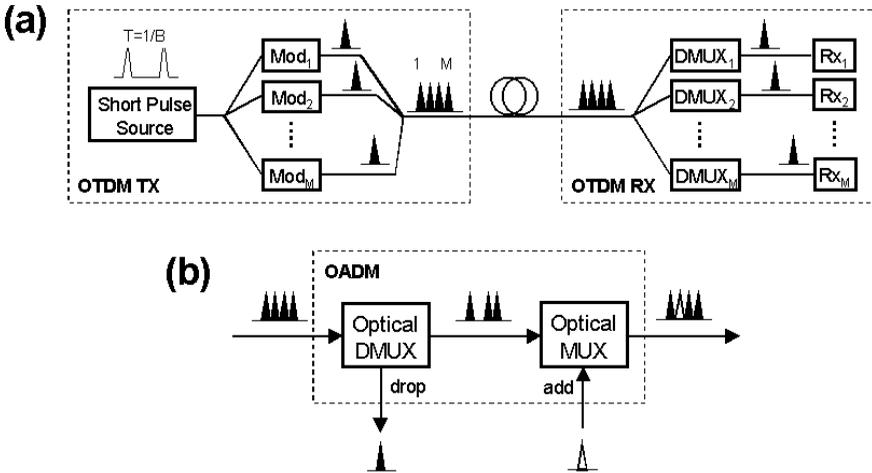


Fig. 17.4. (a) Basic schematic of an OTDM transmitter and receiver in a point-to-point system, (b) concept of OTDM optical add-drop multiplexing

makes photonic integration essential for commercial viability. Compact optical short pulse sources using a tandem of integrated electro-absorption modulators [8] and further integration with semiconductor optical amplifiers (SOA) for compensating losses have been demonstrated [22]. Further chip level integration of splitters, arrays of modulators and amplifiers on PLC platforms also show promise for commercial viability [16]. An OTDM demultiplexing receiver with an integrated electro-absorption modulator for optical gating followed by a detector has been demonstrated at 40 Gb/s [7]. Further integration of optical amplification has also been achieved for better receiver sensitivity performance [21]. Monolithically integrated SOA devices have been used for add-drop applications [28] while a single electro-absorption modulator has also been shown to simultaneously demultiplex and detect a single channel while transparently transmitting the express channels [12]. Using this capability, an alternative OTDM receiver architecture that imitates high-speed electrical receiver sub-systems with the benefit of operating at the tributary base rate has also been proposed [11]. The compact OTDM receiver eliminates the $1:M$ splitter and requires a single input fiber into a series of integrated and cascaded modulators to achieve high-speed serial-optical to direct tributary-speed parallel-electrical conversion. Apart from photonic integration, there are several other challenges that need to be considered for OTDM to have commercial penetration. The generation, multiplexing, and gating of such short pulses imposes very strict and challenging polarization, crosstalk, extinction ratio and timing with active signal processing requirements. Finally, further challenges of optical transmission of ultra-high data rates in dynamic networks will require thoughtful consideration.

17.4 All-Optical Edge Grooming

Present day networks are based on optical circuit switching with dedicated traffic paths carrying traffic from source to destination. As network bandwidth demands increase, DWDM systems with higher count channels and faster data rates are utilized to provide the required bandwidth capacity. While DWDM is a very effective means of using the bandwidth of the installed fiber base, such switching paradigms exhibit poor performance when sub-wavelength traffic is considered. As described in the preceding section, all-optical circuit grooming using OTDM techniques can help in aggregating several optical channels onto one wavelength. However, the limited granularity and static nature of OTDM switching does not render this as an efficient solution in edge networks. Therefore, traffic grooming at the edge of a network, which presently makes use of the fine granularity electronics can offer, is becoming increasingly important to improve the usage of the available optical resources and to increase the overall network utilization. In SONET and SDH networks, this is mainly accomplished by statistical multiplexing, where data from a large number of logical channels are carried on a single wavelength in the

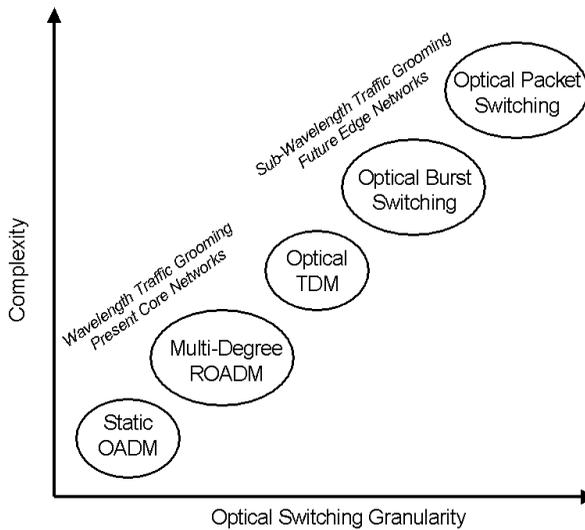


Fig. 17.5. Complexity of the various all-optical traffic grooming technologies

physical medium. The available bandwidth is dynamically allocated only towards active channels, which enables more devices to be connected than other multiplexing techniques. However, as services such as VoD and VoIP are being introduced, the requirements for low latency and high granularity are increasing. Due to the high cost and power consumption of high bit-rate electronics and with progressively more emphasis on transparency, all-optical edge grooming using OPS and OBS are becoming more and more interesting [34].

From a traffic utilization point of view, the highest gains are currently envisaged by using OPS since it can provide both the granularity and latency required for multi-media services [3, 35]. Instead of encapsulating each IP packet into larger SONET frames that are transmitted in a synchronous bit stream, with OPS each IP packet is transmitted by itself with an additional small header. Thus, IP packets arrive at the end destination asynchronously as they are not electronically processed and synchronized at each node in the network. While OPS provides highly efficient and granular switching, it also requires new technology at the edge that also permeates into the whole core network. The current lack of switches with nanosecond switching times and optical buffers are prohibitive factors for OPS networks with substantial development time required.

A hybrid between optical circuit and packet switching is OBS [4]. In contrast to OPS where packets are sent out one by one, OBS networks aggregate packets with the same end destination into larger packets, which are denoted as bursts. With significantly larger packets, the speed requirements of the switching elements in the core network are a lot less strict at the expense of decreased network utilization. The minimum length of a burst is determined

by the switching technology used and has an upper limit for the time allowed in which a burst can be assembled. As a burst cannot arbitrarily wait long enough for IP packets to arrive in assembling the burst, many bursts are likely to be timed out and transmitted without the minimum number of IP packets, in which case the network utilization is decreased. While OBS alleviates the issue of fast switching, it still shares the issue of optical buffers as in OPS.

The complexity of the various technologies discussed in this chapter for all-optical traffic grooming is summarized in Fig. 17.5 with respect to switching granularity.

17.5 Conclusion

In the quest for high-capacity and low-cost optical networks, traditional electronic traffic grooming is being replaced by more efficient all-optical technologies. Multi-degree ROADMs with wavelength grooming capabilities are already being deployed in commercial core networks while active research in OTDM, OPS, and OBS is underway for all-optical sub-wavelength grooming of edge traffic.

References

1. E.B. Basch, R. Egorov, S. Gringeri, and S. Elby. Architectural tradeoffs for reconfigurable dense wavelength-division multiplexing systems. *IEEE Journal in Selected Topics on Quantum Electronics*, 12:615–626, 2006.
2. G. Baxter, S. Frisken, D. Abakoumov, Z. Hao, I. Clarke, A. Bartos, and S. Poole. Highly programmable wavelength selective switch based on liquid crystal on silicon switching elements. In *Optical Fiber Communications Conference (OFC/NFOEC), paper OTuF2*, 2006.
3. D.J. Blumenthal, J.E. Bowers, L. Rao, H.-F. Chou, S. Rangarajan, and H.N. Poulsen. Optical signal processing for optical packet switching networks. *IEEE Communications Magazine*, 41:S23–S29, 2003.
4. Y. Chen, C. Qiao, and X. Yu. Optical burst switching: A new area in optical networking research. *IEEE Network Magazine*, 18(1):16–23, 2004.
5. O. Gerstel and H. Raza. Predeployment of resources in agile photonic networks. *IEEE Journal of Lightwave Technology*, 22:2236–2244, 2004.
6. A.H. Gnauck, G. Charlet, P. Tran, P. Winzer, C. Doerr, J. Centanni, E. Burrows, T. Kawanishi, T. Sakamoto, and K. Higuma. 25.6-Tb/s C+L-band transmission of polarization-multiplexed RZ-DQPSK signals. In *Optical Fiber Communications Conference (OFC/NFOEC), paper PDP19*, 2007.
7. V. Kaman, Y.J. Chiu, T. Liljeberg, S.Z. Zhang, and J.E. Bowers. A compact 40 Gbit/s demultiplexing receiver based on integrated tandem electroabsorption modulators. *IEEE Electronics Letters*, 36:1943–1944, 2000.
8. V. Kaman, Y.J. Chiu, T. Liljeberg, S.Z. Zhang, and J.E. Bowers. Integrated tandem traveling-wave electroabsorption modulators for > 100 Gbit/s OTDM applications. *IEEE Photonics Technology Letters*, 12:1471–1473, 2000.

9. V. Kaman, R.J. Helkey, and J.E. Bowers. Compact and scalable 3-D MEMS optical switches. *Journal of Optical Networks*, 6:19–24, 2007.
10. V. Kaman, R.J. Helkey, and J.E. Bowers. Multi-degree ROADMs with agile add-drop access. In *Photonics in Switching Conference, paper TuA2*, 2007.
11. V. Kaman, A.J. Keating, S.Z. Zhang, and J.E. Bowers. Electroabsorption modulator as a compact OTDM demultiplexing receiver. In *ECOC, paper 9.4.6*, 2000.
12. V. Kaman, A.J. Keating, S.Z. Zhang, and J.E. Bowers. Simultaneous OTDM demultiplexing and detection using an electroabsorption modulator. *IEEE Photonics Technology Letters*, 12:711–713, 2000.
13. V. Kaman, S. Yuan, O. Jerphagnon, R.J. Helkey, and J.E. Bowers. Comparison of wavelength-selective cross-connect architectures for reconfigurable all-optical networks. In *Photonics in Switching Conference, paper 14.1*, 2006.
14. V. Kaman, X. Zheng, S. Yuan, J. Klingshirn, C. Pusarla, R.J. Helkey, O. Jerphagnon, and J.E. Bowers. A 32×10 gbit/s DWDM metropolitan network demonstration using wavelength-selective photonic cross-connects and narrow-band EDFAs. *IEEE Photonics Technology Letters*, 17:1977–1979, 2005.
15. G. Karmous-Edwards. Today's optical network research infrastructures for E-Science applications. In *Optical Fiber Communications Conference (OFC/NFOEC), paper OWU3*, 2006.
16. K. Kato and Y. Tohmori. PLC hybrid integration technology and its application to photonic components. *IEEE Journal in Selected Topics on Quantum Electronics*, 6(1):4–13, 2000.
17. J. Kim, C.J. Nuzman, B. Kumar, D.F. Liewen, J.S. Kraus, A. Weiss, C.P. Lichtenwalner, A.R. Papazian, R.E. Frahm, N.R. Basavanahally, D.A. Ramsey, V.A. Aksyuk, F. Pardo, M.E. Simon, V. Lifton, H.B. Chan, M. Haueis, A. Gasparyan, H.R. Shea, S. Arney, C.A. Bolle, P.R. Kolodner, R. Ryf, D.T. Neilson, and J. V. Gates. 1100×1100 port MEMS-based optical crossconnect with 4-dB maximum loss. *IEEE Photonics Technology Letters*, 15:1537–1539, 2003.
18. J. Kondis, B.A. Scott, A. Ranalli, and R. Lindquist. Liquid crystals in bulk optics-based DWDM optical switches and spectral equalizers. In *IEEE LEOS Annual Meeting*, 1; 292–293, 2001.
19. E. Lach and K. Schuh. Recent advances in ultrahigh bit rate ETDM transmission systems. *IEEE Journal of Lightwave Technology*, 24:4455–4467, 2006.
20. D.M. Marom, D.T. Neilson, D.S. Greywall, C.-S. Pai, N.R. Basavanahally, V.A. Aksyuk, D.O. Lopez, F. Pardo, M.E. Simon, Y. Low, P. Kolodner, and C. A. Bolle. Wavelength-selective $1 \times k$ switches using free-space optics and MEMS micromirrors: Theory, design, and implementation. *IEEE Journal of Lightwave Technology*, 23:1620–1630, 2005.
21. B. Mason, J.M. Geary, J.M. Freund, A. Ougazzaden, C. Lentz, K. Glogovsky, G. Przybylek, L. Peticolas, F. Walters, L. Reynolds, J. Boardman, T. Kercher, M. Rader, D. Monroe, and L. Ketelsen. 40 Gb/s photonic integrated receiver with -17 dBm sensitivity. In *Optical Fiber Communications Conference (OFC), paper FB10*, 2002.
22. B. Mason, A. Ougazzaden, C.W. Lentz, K.G. Glogovsky, C.L. Reynolds, G.J. Przybylek, R.E. Leibenguth, T.L. Kercher, J.W. Boardman, M.T. Rader, J.M. Geary, F.S. Walters, L.J. Peticolas, J.M. Freund, S.N.G. Chu, A. Sirenko, R.J. Jurchenko, M.S. Hybertsen, L.J.P. Ketelsen, and G. Raybon. 40-Gb/s tandem electroabsorption modulator. *IEEE Photonics Technology Letters*, 14(1):27–29, 2002.

23. S. Mechels, L. Muller, D. Morley, and D. Tillett. 1D MEMS-based wavelength switching subsystem. *IEEE Communications Magazine*, 41:88–94, 2003.
24. B. Mikkelsen, G. Raybon, R.J. Essiambre, K. Dreyer, Y. Su, J.E. Johnson, G. Shtengel, A. Bond, D.G. Moodie, and A.D. Ellis. 160 Gbit/s single-channel transmission over 300 km nonzero-dispersion fiber with semiconductor based transmitter and demultiplexer. In *ECOC, paper PD 2-3*, 1999.
25. M. Nakazawa, T. Yamamoto, and K.R. Tamura. 1.28 Tbit/s - 70 km OTDM transmission using third- and fourth-order simultaneous dispersion compensation with a phase modulator. In *ECOC, paper PD 2.6*, 2000.
26. M.J. O'Mahony, C. Politi, D. Klonidis, R. Nejabati, and D. Simeonidou. Future optical networks. *IEEE Journal of Lightwave Technology*, 24:4684–4696, 2006.
27. S. Pau, J. Yu, K. Kojima, N. Chand, and V. Swaminathan. 160-Gb/s all-optical MEMS time-slot switch for OTDM and WDM applications. *IEEE Photonics Technology Letters*, 13:1460–1462, 2001.
28. H.N. Poulsen, K.S. Jepsen, A.T. Clausen, A. Buxens, K.E. Stubkjaeri, R. Hess, M. Dulk, and G. Melchior. Fast optical signal processing in high bit rate OTDM systems. In *IEEE LEOS Annual Meeting*, volume 1, pages 67–68, 1998.
29. G. Raybon, P.J. Winzer, and C.R. Doerr. 1-Tb/s (10×107 Gb/s) electronically multiplexed optical signal generation and WDM transmission. *IEEE Journal of Lightwave Technology*, 25:233–238, 2007.
30. J.-K. Rhee, I. Tomkos, and M.-J. Li. A broadcast-and-select OADM optical network with dedicated optical-channel protection. *IEEE Journal of Lightwave Technology*, 21(1):25–31, 2003.
31. A.A.M. Saleh and J. M. Simmons. Evolution toward the next-generation core optical network. *IEEE Journal of Lightwave Technology*, 24:3303–3321, 2006.
32. R.E. Wagner, J.R. Igel, R. Whitman, M.D. Vaughn, A. B. Ruffin, and S. Bickham. Fiber-based broadband-access deployment in the United States. *IEEE Journal of Lightwave Technology*, 24:4526–4540, 2006.
33. H.G. Weber, R. Ludwig, S. Ferber, C. Schmidt-Langhorst, M. Kroh, V. Marembert, C. Boerner, and C. Schubert. Ultrahigh-speed OTDM-transmission technology. *IEEE Journal of Lightwave Technology*, 24:4616–4627, 2006.
34. S.J.B. Yoo. Optical packet and burst switching technologies for the future photonic internet. *IEEE Journal of Lightwave Technology*, 24:4468–4492, 2006.
35. M. Zirngibl. IRIS: Optical switching technologies for scalable data networks. In *Optical Fiber Communications Conference (OFC/NFOEC), paper OWP1*, 2006.

Index

- admissible solution, 268, 270, 271, 273, 274, 276
- aggregated demand, 257
- all-optical circuit grooming, 284
- all-optical edge grooming, 286
- all-optical grooming, 279
- all-optical wavelength grooming, 280
- amplifier, 124
- arbitrary alternate routing, 208

- bandwidth-blocking probability (BBP), 113
- bi-directional path networks, 69
- bi-directional ring networks, 69
- blocking probability, 193
 - end-to-end, 201
 - fiber route, 204, 206
 - g-link, 195, 197, 201, 204, 206, 208
 - path, 195, 197, 201, 208
 - segment, 204, 206
 - wavelink, 197, 199
 - active call, 207
 - passive call, 207
 - waveroute, 196, 197, 201, 204
- BXC, 254–256, 260, 261, 269, 271, 276

- call
 - active, 206
 - class, 198
 - multi-service, 198
 - passive, 206
 - service, 198
- carried load
 - fiber route, 205

- g-link, 200
 - active call, 208
 - passive call, 208
- path, 200
 - wavelink, 197, 201, 205
 - waveroute, 200, 205
- central office (CO), 107
- class- (n, k) call, 198
- complete lightpath cycle, 131
- complexity of Traffic Grooming, 57
- composite signal, 255
- connected lightpaths, 130
- control plane support
 - ASON, 21
 - mapping GMPLS, 26
 - call control, 31
 - hierarchical routing, 30
 - link resource manager, 28
 - mapping GMPLS, 27
 - model, 22, 26
 - resource discovery, 28
 - routing areas, 32
 - routing database, 28
 - subnetwork point, 26, 27
 - subnetwork point pool, 26, 27
 - UNI/NNI, 25
- GMPLS, 21
 - adaptation capability, 28
 - bi-directional LSP, 31
 - bundling, 28
 - call control, 31
 - exchange point, 25
 - explicit label, 31

- explicit route, 31
- forwarding adjacency, 28, 30
- generalized label, 27, 31
- hierarchical routing, 30
- hierarchy, 22
- label set, 31
- link management, 22, 29
- LSP, 22, 26
- mapping ASON, 26, 27
- MIB, 28
- model, 22, 26
- OPEX reduction, 21
- overlay/peer/augmented, 25
- resource discovery, 28
- routing protocols, 22
- RSVP-TE, 22
- signaling, 30
- signaling protocols, 22
- SRLG, 29
- support for VCAT/LCAS, 32
- switching capability, 27, 28
- TE link, 22, 26
- testbeds, 35
- upstream label, 31
- virtual link, 30
- grooming data plane
 - OTN, 22, 24
 - SONET/SDH, 22, 24
 - VCAT/LCAS, 24, 33, 34
- L1 VPN, 33
- PCE, 30, 34
- SLA/SLS, 33
- cost function, 261, 272
- cost per bit, 256
- crossbar switch, 17

- data centers, 253
- database backups, 253
- dedicated backup multiplexing, 152
- dedicated backup reservation, 150
- dedicated protection, 276
- demand time conflict reduction, 164
- dependable dynamic traffic grooming, 149
- deterministic, 254
- Digital Cross-Connect System (DCS), 15
- distribution
 - exponential, 111
 - Poisson, 111

- Dynamic Bandwidth Allocation (DBA), 51
- dynamic provisioning, 142
- dynamic restoration, 138
- dynamic traffic, 5

- electrical aggregation, 257
- electrical cross-connect, 254, 255
- electrical grooming, 254, 256, 257, 262, 272, 276
- elemental network topologies, 61
- End-to-end call loss probability, 197
- EXC, 254–257, 259, 260, 262, 266, 276

- Fiber Channel (FC), 108
- Fiber Connection (FICON), 108
- Fiber Cross-Connect (FXC), 12
- fixed-point equation, 197, 202

- g-link, 195, 196
 - carried load, 197
 - repeated, 209
 - support, 195
 - topology, 195
- general shared path protection, 144
- Generic Framing Procedure (GFP), 52, 108, 215
- grid networks, 69
- grooming, 254
- grooming and routing (GR), 123
- grooming lightpath, 257–259, 263, 264, 272, 273
- grooming link, 195
- grooming strategy, 256
- grooming switches, 15–19
- grooming, routing and wavelength assignment (GRWA), 124

- heuristics, 254
- hierarchical grooming
 - clustering, 79
 - general topology networks, 77
 - hub selection, 79
 - lightpath grooming, 83
 - logical topology, 81
 - performance results, 83
 - ring networks, 74
 - RWA, 82
 - star networks, 76

- torus networks, 76
- traffic routing, 81
- tree networks, 76
- hierarchical grooming algorithm, 79
- I/O ports, 256, 260
- Integer Linear Programming (ILP), 124
- inter-office (IOF), 107
- inverse multiplexing, 109
- Lagrangian relaxation, 124
- light tree, 18
- light-trail, 155, 165
 - dynamic, 168
 - static, 166
- lightpath, 2, 254–256, 260, 261, 264, 268–270, 272, 273
- lightpath cycle, 131
- lightpath sequence, 194
- Line Terminal Equipment (LTE), 59
- linear programming, 254
- Link Capacity Adjustment Scheme (LCAS), 51, 108
- link protection, 276
- link-based protection, 140
- logical topology, 2, 195
 - dynamic, 195
- management burden, 256
- Manhattan street networks, 69
- many-to-one traffic grooming, 227–232
 - exact approach, 229
 - heuristic approach, 231
- marginal demands, 259, 264
- matrix algebra, 263
- meta-heuristics, 254
- MG-OXC, 255, 260, 261
- mixed primary-backup grooming policy, 149
- monetary cost, 254
- Multi-Commodity Flow (MCF), 64
- multi-domain traffic grooming
 - conclusions, 251
 - introduction, 237
 - multi-domain standards, 238
 - open problems in multi-domain grooming, 245
 - research studies, 242
- multi-granularity switching, 255, 256
- multi-hop full grooming OXC, 17
- multi-hop partial grooming OXC, 18
- multi-layer, 255, 256, 262, 276
- Multi-Service Provisioning Platform (MSPP), 15
- multicast traffic grooming, 214–227
 - dynamic multicast traffic grooming, 222–227
 - enabling technology, 215
 - static multicast traffic grooming, 216–222
 - exact approaches, 216–220
 - heuristic approaches, 220–222
- multipoint traffic grooming, 213–233
 - types, 213
- network cost, 256, 261–263, 272, 276
- network design, 254
- network layer, 256, 261, 276
- node-edge incidence matrix, 126
- nominal capacity, 257
- NP*-Complete, 57, 112, 123
- objective function, 261
- OEO, 3
 - switching cost, 3, 60
- offered load
 - active call, 207
 - g-link, 200
 - passive call, 207
 - segment, 205
 - wavelink, 197, 201, 205
 - waveroute, 197
- optical burst switching (OBS), 287
- Optical Cross-Connect (OXC), 12, 114, 137
- optical grooming, 254, 256, 260–262, 268, 270, 273, 274, 276
- optical layer, 125
- optical node architectures, 9–19
- optical packet switching (OPS), 287
- optical switching, 13
- optical time-division multiplexing (OTDM), 284
 - principle, 285
- Optical Transport Network (OTN), 108
- optimisation techniques, 254
- opto-electronic, 256
- p*-Cycle protection, 141

- partial protection, 153
- partition, 269
- path protection, 276
- path restoration algorithm, 142
- path-based protection, 140
- performance model, 193
- permanent connection, 256
- permanent demands, 254
- physical layer, 125
- planning, 254
- Point-to-Point Protocol (PPP), 108
- power, 269
- pre-cross-connected shared path
 - protection, 145
- preplanned protection, 138
- protection design, 143
- provisioning, 254, 256, 273

- random traffic demands, 253
- random traffic grooming, 209
- random wavelength assignment, 209
- Reconfigurable OADM (ROADM), 280
 - multi-degree, 282
- reduced load approximation, 194, 196
- restricted sharing admission, 196, 198
- Routed Scheduled Band Group, 261, 270, 273, 274
- router, 255
- Routing and Wavelength Assignment (RWA), 3, 4

- scheduled demands, 253, 254, 276
- scheduled lightpath demands, 268
- scheduled services, 160
- SED, 254, 256, 257, 262–265, 267, 272, 273, 276
- segment, 204
- segregated primary-backup grooming
 - policy, 149
- self-routing fabric, 17
- service- k call, 198
- set-up date, 253, 254, 261, 262, 268
- shared backup multiplexing, 151
- shared backup reservation, 150
- shared bus, 16
- shared memory, 17
- shared protection, 276
- shelf, 269
- shortest path, 256, 262, 263, 268, 272

- single-hop grooming OXC, 17
- single-layer, 256
- SLD, 254, 256, 261, 262, 267–270, 273, 274, 276
- sliding scheduled traffic model, 162
- Soft Failure, 51
- solution space, 270
- SONET/SDH, 39
 - frame structures, *see* STS
 - layers, 40
 - Next Generation (NG-SONET/SDH), 47
 - rings
 - protection methods, 43
 - types, 43
 - standard rates, 41
 - Virtual Tributaries (VT), 43
- SONET/SDH ring grooming, 91–105
 - arbitrary traffic, 98–105
 - Bidirectional Line-Switched Ring (BLSR), 91
 - fixed traffic, 92–98
 - algorithms, 96–98
 - bounds, 93–96
 - hubbed rings, 95, 98
 - incremental traffic, 104
 - multi-ring topologies, 97
 - traffic types, 91
 - Unidirectional Path-Switched Ring (UPSR), 91
- source-edge incidence matrix, 126
- spider networks, 69
- star networks, 69
- state-of-the-art, 256
- static traffic grooming with min-max
 - objective, 71
- statistical multiplexing gain (SMG), 12
- STM, 41
- storage-area network, 108
- STS, 41
- subgraph-based routing, 140
- subgraph-based routing strategy, 142
- subwavelength demands, 3, 253–256
- survivability, 138
- switch ports, 254
- switching fabric, 255, 256
- switching granularity, 253

- tear-down date, 253, 254, 261, 262, 268, 276
- time diagram, 257, 272
- time-disjoint, 254, 276
- topology, 254, 257, 260–263
- traffic grooming, 159
 - dynamic, 193
 - multi-hop, 194
 - policy, 206
 - single-hop, 194
 - static, 193
- traffic pattern, 111
- transit traffic, 256
- transponder, 12, 124
- transponders, 254
- tree networks, 71
- tuple, 254, 261

- uncertainty, 254, 276
- unidirectional path networks, 62
- unidirectional ring networks, 69

- ventilation system, 269
- Virtual Concatenation (VC/VCAT), 48, 108, 109
- Virtual concatenation group (VCG), 48
- virtual topology, 2

- waveband, 253–256, 260, 261, 270, 271
- waveband cross-connect, 254, 255
- waveband switching, 261
- waveband-switching connection, 260, 261, 267, 268, 271
- wavelength, 255
- wavelength assignment, 260
- wavelength assignment (WA), 123
- wavelength capacity, 125
- wavelength continuity, 126
- wavelength continuity constraint, 260
- wavelength conversion, 204, 260
- wavelength cross-connect, 254, 255
- Wavelength Interchanging Cross-Connect (WIXC), 13
- Wavelength Selective Cross-Connect (WSXC), 13
- Wavelength-Band Cross-Connect (WBXC), 12
- wavelink, 195, 196, 198
 - repeated, 204, 209
- wavelink blocking model, 198
- waveroute, 195, 196, 198, 204
- WDM grooming, 193
- WXC, 254–256, 259–262, 266, 269, 271, 276

Part III

Frontiers

Multi-Domain Traffic Grooming

Nasir Ghani, Qing Liu, Nageswara S. V. Rao, Ashwin Gumaste

15.1 Introduction

The last decade has seen many advances in high-speed networks. At the fiber level, dense wavelength division multiplexing (DWDM, Layer 1) has gained favor as a terabit solution, capable of “lightpath” circuit routing. Meanwhile next-generation SONET/SDH (NGS) (Layer 1.5) [1] has gained strong traction in the metro/edge, providing flexible “sub-rate” aggregation. Finally, Ethernet and IP networks (Layers 2, 3) have seen new quality of service (QoS) provisions via the differentiated services (Diff-Serv) and integrated services (Intserv) frameworks. Related control standards have also emerged, most notably multi-protocol label switching (MPLS) for Layer 2/3 flow-based QoS support. Further generalizations have even adapted this solution for “non-packet-switching” layers, i.e., generalized MPLS (GMPLS) [2].

As the above solutions undergo widespread deployment, a complex interconnection of different data-planes has resulted. From a theoretical aspect, this is embodied by an interconnection of multiple horizontal domains (i.e., networks) comprising of different vertical layers (i.e., technologies), as shown in Fig. 15.1 [3]. These segmentations are based upon various factors, such as technology, scalability, geographic, economic, administrative, etc. Moreover as the scale and reach of services expands, there is a pressing for “end-to-end” provisioning across heterogeneous domains, i.e., horizontal-vertical control. A good example is the growth in “e-science” applications [3]. However, inter-layer provisioning today is still done via manual provisioning of domain-specific control planes, giving high inefficiency and long lead times, i.e., hours to days.

Multi-domain networking has long been supported in IP and ATM networks. Moreover, new efforts are also introducing these capabilities in optical transport networks. Nevertheless, the broader topic of provisioning across heterogeneous network layers has not been addressed in detail. Note that there is a clear distributed multi-domain grooming aspect to this problem as different circuit/flow granularities are involved, i.e., gigabit wavelengths, “sub-rate”

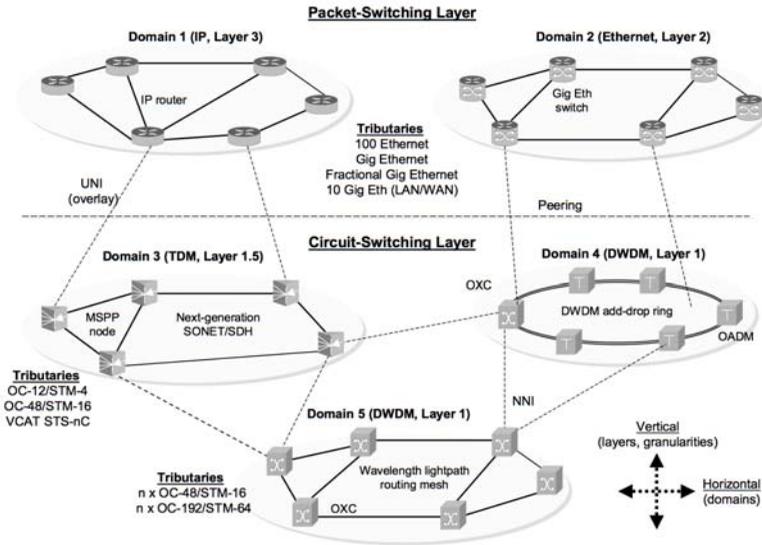


Fig. 15.1. Multi-layer, multi-domain networking infrastructures

SONET/SDH tributaries, MPLS label switched paths (LSP), etc. Overall, this is a very difficult problem owing to some key complicating factors. Foremost, there are no unified control plane standards for integrating multiple layers. In addition related provisioning algorithms are lacking, particularly those operating in distributed (intra, inter-carrier) settings with no/partial global state.

This chapter addresses the emerging area of multi-domain grooming and is organized as follows. First, Section 15.2 presents a brief survey of standards for multi-domain networks. Subsequently Section 15.3 surveys some work on multi-domain protocols and algorithms. Section 15.4 then delves into the relatively unexplored area of multi-domain grooming and details key directions in inter-layer routing, path computation, signaling, and survivability. Finally Section 5 presents overall conclusions and future directions.

15.2 Multi-Domain Standards

Over the years, a wide range of multi-domain networking standards have been developed by various standards organizations, including the IETF, ITU-T, and OIF. These efforts are briefly reviewed here.

15.2.1 ITU-T Standards

The ITU-T has matured a comprehensive automatically switched transport networks (ASTN) framework for multi-domain optical networks. The reference architecture for ASTN is G.8080 [2] and describes a group of components to control transport resources to setup, maintain, and release a client layer connections. However, the internal topology and connectivity of the underlying optical layers is not made visible to the client layer and is instead treated as a sub-network point pool (SNPP) link (“virtual” link). Now ASTN follows a hierarchical design in which designated entities perform control for single and multiple layers. For example, its routing hierarchy consists of areas with requirements for inter-area auto-discovery, auto-provisioning and auto-restoration. However, this model is flexible and each domain’s internal control plane can be tailored to the particular types and capabilities of the equipment within the domain. Nevertheless, this overall framework only addresses architectural design and no explicit protocols specifications are provided. It is here that developments within the IETF and OIF are of crucial importance.

15.2.2 OIF Standards

The OIF has developed several control standards for optical network interfacing, including a client–carrier user-network interface (UNI) and a carrier–carrier network node interface (NNI). The UNI implements bandwidth signaling for client devices (i.e., IP/MPLS routers) to request/release “optical” connections from underlying carrier SONET/SDH or DWDM domains, i.e., “optical dialtone”. Since there is no trust relationship here, carrier topology information is not propagated to the client side, i.e., overlay model [2]. The latest UNI 2.0 features much improved capabilities for security, bandwidth modification, etc. Meanwhile, the NNI interface implements inter-domain functions for reachability/resource exchange and setup signaling. Now two variants have been designed here, interior NNI (I-NNI) and external-NNI (E-NNI). The former interfaces nodes within the same administrative area whereas the latter serves adjacent areas. E-NNI relegates all interfacing issues to domain boundaries, thereby removing restrictions on domain-internal control and equipment interoperability. Furthermore, NNI adopts a link state routing approach, e.g., via single/stacked instances of GMPLS routing protocols.

The overall UNI–NNI combination facilitates rapid deployment of new services across multi-vendor “optical” domains. Moreover, this framework works for both optical DWDM and SONET/SDH domains, as facilitated by the different granularities supported in UNI/NNI signaling. Hence the NNI standard can support multi-layer interfacing at the circuit-switching level. However, since there is no resource exchange across the UNI, higher layer IP/MPLS networks must treat underlying “optical” links as tunneled links. This setup will be problematic for client layer routing since the number of Layer 3 adjacencies will grow per the square of border nodes, i.e., connection

mesh problem. Clearly, as the number of border nodes grows, this routing setup becomes less scalable, considering the amount of adjacency information has to be maintained. Hence clients may have to run data-plane inter-domain routing protocols across UNI-provisioned connection links.

15.2.3 IETF Standards

Well-established IP network architectures comprise a hierarchy of autonomous systems (AS) and routing areas (domains). Within areas, routers run associated interior gateway protocols (IGP) such as open shortest path first (OSPF) or intermediate-system to intermediate-system (IS-IS) [2]. These protocols use link-state routing to maintain link state databases (LSDB). Meanwhile at the inter-AS level, commensurate exterior gateway protocols (EGP) are used. The mainstay here is border gateway protocol (BGP) which runs between exterior gateway nodes and uses distance vector routing to provide end-point reachability exchange. In addition, OSPF also provides an additional level for disseminating “higher-layer” aggregated state between area border routers (ABR). Note that many operators will want to limit internal state dissemination in inter-carrier settings owing to privacy and security concerns.

With growing flow-level QoS requirements in IP networks, new extensions have also been proposed for OSPF. For example, OSPF-traffic engineering (OSPF-TE, RFC 2676) provides extensions and opaque link state (LSA) definitions for “QoS-related” link state to support advanced constraint-based routing (CBR). Meanwhile there have also been proposals for augmenting BGP messaging with added QoS attributes. For example, [4] exchanges QoS-enabled reachability information via new BGP message attributes. However, these offerings do not alter the inherent distance vector design. Finally, the hierarchical routing has also been implemented in older asynchronous transfer mode (ATM) networks via the private network-to-network interface (PNNI) protocol, i.e., via peer group leaders.

Meanwhile the GMPLS framework devises new routing/signaling versions to manage “circuit-switched” entities, e.g., TDM circuits and DWDM light-paths [2]. For example, new OSPF-TE opaque LSA definitions have been added for DWDM and SONET/SDH links. Hence TE databases (TEDB) can now store information fields for wavelengths/usages, timeslots/usages, shared risk link groups (SRLG) diversity, etc. New enhancements have also been added to the reservation protocol, RSVP-TE, to implement DWDM and SONET/SDH circuit setup/takedown, e.g., hard state, recovery, etc. Additionally, RSVP-TE also provides a loose route (LR) feature to specify a partial (high-level) node sequences along with signaling-based explicit route (ER) expansion to resolve the full path route. Finally a new link management protocol (LMP) has been developed for fault discovery and localization.

However, GMPLS does not provide any specific routing extensions for multi-domain BGP as its inherent distance-vector design is not well-suited

for circuit routing. Therefore, two-level OSPF-TE is the most germane inter-domain protocol for GMPLS. Nevertheless, OSPF-TE assumes a peer-to-peer model in which all nodes perform identical routing/signaling. Clearly, today's emerging diversified, multi-layer domains settings will prove problematic here. For example, many vendor solutions may not provide GMPLS support and instead use centralized vendor-proprietary network management systems (NMS) or operations support systems (OSS) and TL-1 messaging [1]. In these cases, proxy solutions can be used to translate external vendor-specific control formats [5].

The IETF is also addressing TE path computation via its path computation element (PCE) architecture [6, 7]. This framework allows path computation clients (PCC) to interact with PCE entities to resolve constraint-based routes. A PCE either resides in a standalone manner or is co-located with a network node and has access to domain-level resource/policy databases. This architecture provides a request/response protocol for PCC-PCE and PCE-PCE interaction, i.e., PCE protocol (PCEP) [7]. Finally, automated discovery allows PCC entities to locate a PCE with requisite capabilities, e.g., in terms of network layers, algorithms, backup path computation, etc. Two PCE computation models have also been proposed, centralized, and distributed [6]. In the former, all path computation is performed by a single PCE entity. Although this is realistic for intra-domain scenarios, the broader multi-domain case necessitates this PCE to have global resource and policy visibility. Clearly, a single entity also poses notable scalability and reliability limitations.

Meanwhile, distributed computation uses multiple PCE entities to resolve end-to-end paths. This approach also supports "domain-specific" constraints, such as bandwidth/delay in IP domains and wavelength/color continuity in DWDM networks. Moreover, distributed PCE computation is very germane for multi-domain/multi-layer/multi-carrier settings with partial visibility. Herein, two approaches are outlined to handle the contingencies of varying levels of "global" state, i.e., multi-PCE path computation with and without inter-PCE signaling [7]. Overall both of these schemes induce close couplings between inter-domain routing and signaling.

Multi-PCE computation without inter-PCE signaling requires the head-end PCE (i.e., signaled by PCC) to compute a partial or loose route to the destination. This route can be a "domain-level" sequence of border gateways which is then inserted into appropriate downstream signaling messages. Here receiving entities (border gateways) consult their own PCE entities to "expand" the strict-hop sequences across their domains to the next partial/loose hop. This "inter-PCE" approach fits in well with the above-mentioned RSVP-TE LR/ER features. Meanwhile, multi-PCE computation with inter-PCE signaling requires the head-end PCE to recursively request other PCE entities to compute sub-path segments. This setup usually returns a complete strict-hop source-destination route to the requesting PCC and is more suitable for little/no inter-domain visibility. Alternatively, in inter-

carrier settings, this approach may still return a loose route sequence, thereby preserving intra-domain route confidentiality, see [7].

15.3 Research Studies

Many multi-domain research studies have been conducted. However, most of these efforts have focused on single-layer (homogeneous) packet/cell-switching networks, Layers 2–3. Recently some work on multi-domain optical networks (Layer 1) has also emerged. As yet, very few studies have looked at multi-domain/multi-layer inter-networking.

15.3.1 Multi-Domain Packet/Cell-Switching Networks

Many studies have looked at multi-domain ATM (Layer 2) and IP (Layer 3) networks. In particular, graph-based topology abstraction schemes for state reduction [8–11] have been a key focus area. These strategies transform a domain topology into a reduced “virtual” graph with fewer abstract vertices and edges. This is typically done by a designated domain entity, e.g., routing area leader (RAL) [2], which then propagates the abstract link state to other gateway nodes to maintain “globalized” state. Expectedly, the computational entity must have access to intra-domain state. Overall, topology abstraction allows operators to control internal state dissemination and effectively couples intra- and inter-domain routing.

An early study of abstraction in hierarchical ATM PNNI networks is presented in [8]. This effort details algorithms for computing complex node representations to minimize the number of edges/links in a peer group summarization. The findings here show much-reduced computation times and over an order magnitude reduction in domain state. Also, [9] extends abstraction to multi-domain IP QoS networks, summarizing bandwidth state using various topologies, i.e., star, mesh, tree, and spanner graphs. These reductions are tested with various path computation strategies, including widest shortest, shortest-distance, etc. The overall findings confirm improvements in routing scalability and reduced routing fluctuation. Meanwhile, [10] looks at topology aggregation in directed graphs with additive link metrics (delay, cost). This problem is treated from an information-theoretic perspective and related bounds for compression distortion are derived based upon the asymmetry constant. Furthermore, [11] develops schemes to incorporate both bandwidth and delay parameters via line segmentation techniques. Various Dijkstra-based heuristics are then used to compute paths meeting dual bandwidth and delay constraints. Overall, the results show improved success rates and lower crankback messaging loads.

Finally, studies have also looked at QoS enhancements for BGP distance vector routing. For example, [4] proposes messaging extensions to convey

additional QoS state for paths to destination domains/prefixes. This is coupled with pre-engineered QoS-based service level specifications (SLS) between providers along with “meta-QoS class” planes extending across multi-carrier AS. Results for sample Internet topologies show best performance when paths selection is done using both bandwidth and delay constraints. In addition, messaging overheads are shown to be minimal.

15.3.2 Multi-Domain Optical Networks

Multi-domain DWDM networks are also becoming a key focus. To date, most efforts in this area have focused on distributed lightpath RWA and signaling in all-optical and opto-electronic networks. For example, [12] details a domain-by-domain scheme in which gateways maintain complete (alternate) route state across all-optical and opto-electronic networks. Detailed simulation shows the overall effectiveness of the proposed model. However, this setup is more favorable to BGP-type implementations and related resource propagation (path dissemination) results are not shown.

Meanwhile, [13] studies DWDM networks comprising of multiple segments, i.e., domains. Here, three different inter-domain wavelength routing algorithms are developed, i.e., end-to-end (E2E), concatenated shortest path (CSR), and hierarchical routing (HIR). The E2E basically assumes a “flat” globalized graph whereas HIR assumes a hierarchical graph with segments summarized as nodes. Meanwhile CSR simply uses local information to perform segment-by-segment routing. Numerical results with a specialized mesh-torus topology show significant blocking reduction with the E2E scheme for “multi-segment” requests. However, the CSR scheme does not provide any intra-domain state and further inter-domain routing algorithms are not addressed.

Now, very few studies have looked at routing in multi-domain DWDM networks. Here, it is generally very difficult to re-apply IP/MPLS abstraction schemes since resource constraints are different. Hence recent studies have looked at new DWDM-based abstraction schemes. For example, [14] presents a theoretical treatment of abstraction with border node wavelength conversion. Here, various information models are developed and lightpath selection is modeled as a Bayesian decision problem. The findings for bus topologies show that scalable information models can achieve a good trade-off between performance loss and the amount of network state. Nevertheless, inter-domain routing and RWA issues are not addressed.

Meanwhile, [15] tables a hierarchical inter-domain solution using simple-node abstraction. Nevertheless, signaling provisions are not considered here and only the all-optical case is handled. Conversely, [16] presents a much more detailed study of abstraction in multi-domain all-optical/opto-electronic DWDM networks. Namely, graph theoretic RWA schemes (via k-shortest path heuristics) are used to generate full mesh and star abstractions. Furthermore, detailed inter-domain link-state routing protocols and distributed lightpath

RWA algorithms are also specified (using OSPF-TE, RSVP-TE). Findings show very good blocking reduction and lower inter-domain signaling loads.

Dynamic sub-path protection for DWDM networks has also been considered. For example, [17] segments end-to-end lightpaths into independent “protection domains” and implements shared segment protection with/without wavelength conversion. The algorithm pre-defines self-healing cycles in the network topology and attempts to allocate self-healing loops for lightpaths accordingly. Overall, it is shown that maximum sharing can be achieved by properly assigning wavelengths to the shared channels. Nevertheless, these schemes assume full global state, i.e., “flat/single-domain” network, and hence their application in distributed multi-layer settings is not straightforward. Finally, Layer 1 virtual private networks (Layer 1 VPN) provisioning in multi-domain SONET/SDH network has also been studied in [18]. Layer 1 VPN represents a new service model in which transport-layer resources are partitioned to build multiple “virtual infrastructures” over a single physical network. Here, the authors develop a novel OSPF-TE link model to share time-slots between client VPN topologies. Furthermore, non-adjacent L1 VPN clients are interconnected via “virtual link” multi-domain SONET/SDH connections. Overall, results show good sharing efficiency and blocking reductions with shortest-widest path selection. However, all virtual links here are pre-specified (static), as per end-user demands, and inter-domain routing is not done.

15.3.3 Multi-Domain/Multi-Layer Networks

As mentioned in Section 15.1, there is a clear grooming aspect to multi-domain/multi-layer networks. For example, multiple fine-granularity label switched paths (LSP) originating in Layer 3 domains can be aggregated over coarser DWDM lightpaths. In a similar manner, sub-rate SONET/SDH circuits can also be groomed over DWDM lightpaths. Now, even though grooming algorithms have been widely studied over the years, most studies have focused on idealized settings with full resource state knowledge across all layers. The broader topic of distributed multi-domain grooming, i.e., in the presence of limited global state, is largely unexplored and only a handful of studies have been conducted to date.

In particular, [19] presents one of the first studies on distributed SONET-DWDM provisioning in multi-domain networks. Here, the authors define a multi-segment graph model (with boundary grooming) and evaluate various path selection schemes, e.g., centralized (full-knowledge), domain-by-domain (local knowledge), and hierarchical source routing (partial inter-domain knowledge). The latter approach only propagates domain-internal state for a specified granularity levels, although detailed specifications for state compression are not given. Overall, results show much-improved blocking performance with increasing levels of inter-domain state. Furthermore, [20] studies grooming in SONET-DWDM networks and uses threshold-based triggers to add/delete lightpaths. Here, both centralized (global) and distributed schemes

are developed, with the latter only using traversing path state at a node, i.e., no routing. Overall, results show good performances, contingent to grooming ports and request granularities. However, no inter-domain considerations are discussed.

Finally, [21] tables a novel framework for direct Ethernet-over-DWDM integration. This scheme devises an integrated signaling-based approach to simultaneously provision full wavelength and fine granularity Ethernet virtual connections (EVC) at the optical layer. In contrast to conventional provisioning, signaling messages for sub-lambda EVC requests are only processed at the source node of the lightpath. Nevertheless, the authors assume that the underlying DWDM network is relatively static and hence routing is only done at the Ethernet layer. Detailed performance results are presented for various routing update strategies using crankback. Overall, results show minimal increase in setup latencies and signaling loads.

15.4 Open Problems in Multi-domain Grooming

Despite some standards progress, the study of multi-domain/multi-layer grooming remains in its infancy. Now, it is clear that the scope/scale of the problem necessitates distributed, decentralized solutions here. Along these lines, some key open problems need to be addressed, i.e., multi-domain routing, constraint-based path computation/signaling, and survivability.

15.4.1 Multi-layer Routing and State Exchange

Multi-domain routing is a very challenging area, especially in the presence of multiple layers. The most notable issue is how to propagate different types of link or domain state in a such way that a scalable and secure routing framework can be achieved. By and large, multi-domain grooming will benefit from the availability of “global” inter-domain state. Moreover, as per Sections 15.2.3 and 15.3.1, link-state routing is the most suitable scheme for constraint-based provisioning of QoS LSP or TDM/DWDM circuit paths. Hence, it is very desirable to implement some form of hierarchical routing at select border gateways to disseminate various types of link state (as shown in Fig. 15.2):

- **Physical inter-domain links:** These links interconnect nodes in different domains and can span a full range of types, e.g., Gigabit/10 Gigabit Ethernet, SONET/SDH OC-n links, DWDM links, etc. Hence, related multi-layer TEDB must maintain state for all of these link types either via running multiple routing protocol instances or by using a single unified routing protocol. For example, two-level OSPF-TE already provides LSA definitions for both packet-switching links and “non-packet” TDM and DWDM links.

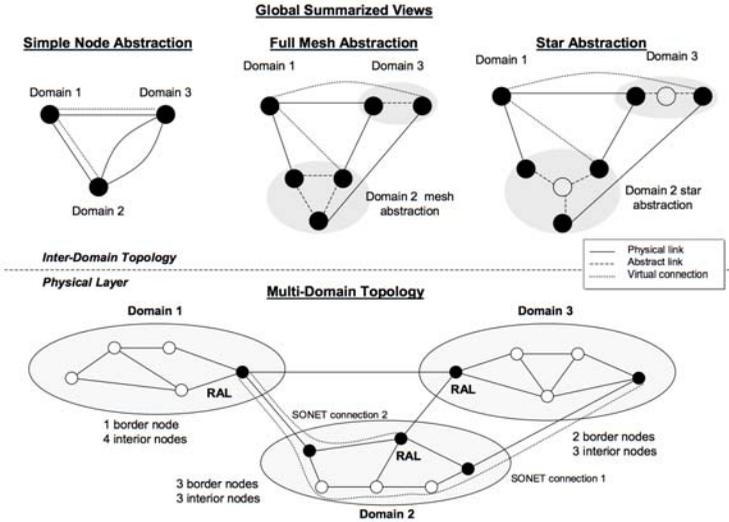


Fig. 15.2. Topology abstraction and link types

- Abstract links:** Abstract links are computed entities used to summarize “layer-specific” domain-level resources and do not reflect physical elements [9]. Although abstraction has been well-studied for IP networks (Section 3.1), circuit-switching networks are more specialized, see initial work in [16]. Hence, further considerations are needed for SONET/SDH grooming links, VCAT features, inverse multiplexing, etc. In addition, new abstractions must capture survivability-related information via abstract links, e.g., fiber/span protection, diversity/risk groups, shared/dedicated resources, etc. In general, expanded attributes can also be developed as these link types need not be constrained to physical link types (in respective domains).
- Virtual link connection:** In general, lower level SONET/SDH and DWDM connections/sub-connections will appear as virtual links to higher-layer devices (i.e., IP/MPLS, Ethernet nodes). These entities can thus be treated as logical TE link (adjacencies) between the two end-points, see Fig. 15.2. Therefore, it will be beneficial to also advertise the resource levels for these entities at the inter-domain level as well, since doing so can yield very good gains in grooming efficiencies. However, there is a clear scalability problem here as the number of such connections can grow in the order of the square of the number of border nodes, i.e., the connection mesh problem as discussed in Section 15.2.2. Hence, a key challenge here will be to develop link aggregation techniques to effectively compress virtual connection link state.

In light of the above, it is also important to consider associated update triggering strategies. To date, various policies have been developed here to achieve update frequency reduction/scalability, e.g., timer-based, absolute change, relative change, and hysteresis-based [20]. All schemes here—except for the former type—are threshold-based. Extensive results for single-domain networks show that relative updates are generally most effective in lowering routing loads and ensuring rapid change sensitivity [20]. Now physical inter-domain link updates can be treated in much the same manner as intra-domain links. Namely, if the underlying resources (bandwidth, timeslots, wavelengths) change as per the chosen metric, an update can be flooded by the sourcing gateway. However, careful analysis is needed to delimit related inter-domain holdoff timers (IHT), which will tend to be larger than corresponding intra-domain values. Virtual link connections can also be treated as such since they are directly associated with underlying connections. Nevertheless, since the number of such connections can be large, further research into scalable triggering policies is needed. Meanwhile, triggering policies for abstract links are more complicated as they pertain to “non-existent” entities. Here, various policies have been developed for abstract IP and DWDM links, see [9, 16], and these can be adapted further. For example, a basic approach is to compute abstractions on a fixed time interval. However, this scheme yields excessive messaging/computational loads for small update intervals and inaccurate state for large update intervals. It is here that relative change measures can give much better performance. For example, [16] uses periodic abstractions but only sends updates if sufficient relative change occurs.

15.4.2 Multi-Layer Grooming and Signaling

Path computation implements the core of the multi-domain grooming function. As opposed to single-domain settings, full topology information in multi-domain networks is generally lacking and path computation has to be done using aggregated (i.e., partial, dated) topology and resource state. Moreover, the existence of multiple layers implies an inherent grooming component, again differing notably from single domain settings which tend to operate at the same layer, e.g., either packet, timeslot, wavelength. Overall, these aspects make multi-layer grooming involving a very challenging problem. In the end, however, the goal here is to resolve and set up full end-to-end routes in accordance with desired constraints and leverage existing standards, particularly GMPLS. Now owing to the distributed nature of the problem it is evident that setup signaling will be required. Although many specific schemes can be envisioned here, few have been analyzed in detail. Hence two broad strategies are outlined here, i.e., hierarchical and per-domain.

Hierarchical Grooming

Hierarchical grooming uses the inter-domain TEDB state to compute “skeleton” routes from the source domain to the destination domain. For

most practical purposes, this path constitutes the end-to-end border node sequence. This information is then coupled with end-to-end signaling mechanisms (RSVP-TE) to expand the full node sequence, i.e., multi-PCE path computation without inter-PCE signaling [7]. Alternatively, this information can also be coupled with PCEP signaling to fill in an exact node sequence, i.e., multi-PCE path computation with inter-PCE signaling [7].

Now consider the actual computation of “skeleton” routes. This is a very complex problem owing to the multiple resource layers involved and the presence of partial/abstracted inter-domain state. Hence, two key areas need detailed investigation here. Foremost, the computing entity (e.g., source domain PCE) must adequately “transform” multi-layer TEDB state to build a “global” multi-layer (granularity) view of the network. The most expedient approach here is to use graph transformations to summarize all link types—physical, abstract, and virtual (Section 15.4.1). This topic is largely unexplored, although augmented graph models have been developed in [9] for multi-layer SONET-DWDM grooming networks (full global state assumption). Clearly, new schemes are needed for adding edges/vertices to capture virtual link connections as well. Next, modified shortest-path schemes need to be studied to compute constrained “skeleton” paths over these transformed graphs. It is here that the grooming aspect can be considered, particularly when “higher-layer” sub-rate connections are to be routed over lower-layer circuits (virtual link connections).

Most regular (non-multi-domain) grooming studies use a “two-step” computation approach. Namely, setup is first attempted over existing higher-layer “virtual” links and pending failure, re-attempted at lower transport layers. Overall, this approach yields very good resource efficiency/packing and is equally relevant in inter-domain scenarios. For example, the initial attempt can search for a “skeleton” route on only those links in the augmented graph (physical inter-domain, abstract, virtual connection) which run at the same granularity/layer as the sourcing node. Here various graph theoretic schemes can be developed to achieve trade-offs between objectives such as hop count, load balancing, minimum cost, e.g., see studies on widest-shortest, shortest-widest, and minimum cost [9].

If a feasible “skeleton” path is not found, further attempts can try to set up new underlying virtual link connection between one/more border gateways. However, this step opens up a whole new dimension that is not present in single-layer multi-domain settings. Namely, the main challenge is to select the actual border node pair(s) between which to initiate virtual connection link(s). This concern has not been addressed in most grooming studies which exclusively assume that all higher layer nodes are directly connected to underlying transport nodes. Clearly this is not the case in general settings in which interior IP/MPLS nodes will not be directly connected to lower-layer SONET/SDH or DWDM transport nodes. One possibility is to use exhaustive search algorithms to enumerate multiple virtual links combinations, e.g., as

done in [16]. Another simpler possibility may be to pre-engineer static virtual connection links using offline optimizations.

Upon computation of a feasible “skeleton” path, appropriate setup signaling schemes are needed as well. In general, these schemes can leverage the ubiquitous RSVP-TE protocol. Specifically, resources can be reserved on physical inter-domain and “tunneled” virtual connection links and path sequences “expanded” for loose route segments. With regards to the latter, border gateways receiving a setup request will basically run intra-domain algorithms to route requests across their domains. Now if the incoming request is of the same granularity, the border node can compute (or query PCE to compute) a regular traversing as per path constraints. However, in the more general case where the border node represents a lower layer, further parameter translation and intra-domain grooming will be needed, and a wide range of existing (intra-domain) grooming schemes can be re-applied. Furthermore, new RSVP-TE proposals are also tabling LSP “stitching” features to inter-connect diverse end-to-end “segments” (layer paths) into a single connection. Again, these can be leveraged as well.

Overall, hierarchical computation can achieve some level of path “optimality” across domains, i.e., contingent to level of inter-domain state. An “optimal” path here is defined as that computed in the absence of domain partitioning, i.e., for “flat” topology [7]. However, detailed studies are really needed to assess the impact of abstracted state on such “optimality”.

Per-Domain Grooming

Per-domain distributed grooming is ideal for settings with minimal inter-domain visibility. Namely, end-to-end routes are expanded in a domain-by-domain manner with each domain specifying its “next-hop” along with the complete internal route via an egress gateway. Here, intra-domain path expansion will be largely similar to the above-detailed hierarchical approach. However at the inter-domain level, the main challenge is how to specify the next domain. If no inter-domain state is available, the only solution may be to make fixed choices, i.e., pre-specified or by consulting BGP tables [7]. Alternatively, it may be much more beneficial to incorporate “global” state to dynamically the downstream domains, if such state is available. Nevertheless, the detailed evaluation of these schemes in mixed-layer networks has not been conducted. In general, per-domain path computation (without inter-domain state) will suffer from higher blocking and lower resource efficiencies as compared to hierarchical computation. Hence, resultant routes will likely be feasible and opposed to “optimal”.

Note that per-domain grooming performance can be improved by leveraging inter-domain crankback signaling. Although [22] studies this approach in SONET-DWDM networks, more investigations are needed. Specifically, commensurate schemes can be designed using intermediate domain or full end-to-end crankback strategies. Furthermore, it may be beneficial to also include active path state in crankback messages in order to improve the

re-try process. Although multi-layer crankback can yield improved resource efficiencies, it poses key trade-offs in inter-layer control plane complexity. Namely, this solution may lower inter-layer routing overheads (e.g., versus hierarchical computation) but at the same time will likely increase inter-layer signaling and setup latencies. These issues need to be closely studied in order to determine the most amenable grooming strategies, i.e., as per connection granularities, arrival rates, holding times, etc.

15.4.3 Survivability

With limited or no visibility into domain-level state, setting up survivable connections is an extremely complicated issue. Hence, the simple approach of computing diverse routes by source nodes (with partial global state) can only guarantee diversity at the physical inter-domain link level. Clearly, further provisions are needed to avoid link overlap during subsequent intra-domain explicit route expansion even with disjoint abstract links computed. Overall, survivable grooming is a relatively new topic area. For example, [23] has recently proposed some protection schemes for SONET-DWDM grooming networks, e.g., protection-at-lightpath (PAL) and protection-at-connection (PAC). Nevertheless, extending these concepts into multi-domain networks lacking full resource diversity state is very challenging. Furthermore, it is also very desirable to extend some form of multi-tiered survivability support as well, e.g., dedicated, shared, non-protected, etc. To address this topic, two survivability strategies can be considered, protection and restoration.

Protection schemes pre-compute diverse primary/backup routes and are well suited for stringent services. Now in the context of multi-domain/multi-layer networks, a simple approach can be to apply SONET strategies in which dual/multi-homed inter-domain links are coupled with robust intra-domain protection schemes. However, these setups are very costly and resource inefficient. Hence, more elaborate end-to-end domain disjoint techniques can be devised, which are generally more amenable to hierarchical grooming with inter-domain state visibility. For example, loose routing algorithms can be extended to compute domain disjoint primary/back “skeleton” paths over augmented inter-layer graphs. Associated setup (loose route expansion) sequences can then be signaled in parallel. Alternatively, the domain-disjoint requirement can be relaxed to allow sharing of abstract links between primary/backup routes. However, the latter approach mandates complicated signaling provisions to ensure intra-domain disjointness in common domains. Again, many of these concerns are open issues and the combined impact of intra/inter-domain schemes on end-to-end recovery still needs to be addressed.

Conversely protection can also be incorporated with the per-domain grooming approach of Section 15.4.2. For example, primary/backup paths can be signaled in a two-step sequential manner where the working route is used to avoid backup link overlaps. Obviously, this approach will have higher setup delays and will be susceptible to “trap” topologies [7]. To address these

concerns, modified crankback schemes can be considered to achieve protection path diversity with parallel signaling, e.g., as tabled in RSVP-TE extensions for record, exclude and associated route objects (RRO, XRO, ARO) [24]. To date, however, most of these concepts have only appeared various standards draft submissions (MPLS-level only), and more detailed analyses and refinements are indeed lacking.

Finally, multi-domain restoration can also be done using post-fault crank-back recovery and this is most germane for connections with lower latency/guarantee stringencies. Along these lines, various signaling provisions need to be devised, e.g., such as proposed RSVP-TE XRO extensions. However, further studies are needed to analyze the trade-offs between resource efficiencies and signaling loads and also compare versus pre-provisioned protection.

15.5 Conclusion

This chapter studies multi-domain grooming in heterogeneous layered networks, a challenging and largely unexplored problem. Along these lines related standards and research studies are surveyed and key open research problems identified. In particular, the latter issues pertain to multi-layer/multi-domain routing, path computation and signaling, and survivability. Finally, future studies can delve into even more advanced multi-domain problems, e.g., VPN provisioning, advance reservation, etc.

References

1. N. Ghani, et al., "Metropolitan Optical Networks," *Optical Fiber Telecommunications IV*, Academic Press, pp. 329–403, March 2002.
2. G. Bernstein, B. Rajagopalan, D. Saha, *Optical Network Control-Architecture, Protocols and Standards*, Addison Wesley, Boston 2003.
3. N. Rao, et al., "Ultra Science Net: Network Testbed for Large-Scale Science Applications," *IEEE Communications Magazine*, Vol. 3, No. 4, pp. S12–S17, November 2005.
4. D. Griffin, et al., "Interdomain Routing Through QoS-Class Planes," *IEEE Communications Magazine*, Vol. 45, No. 2, pp. 88–95, February 2007.
5. T. Lehman, et al., "Control Plane Architecture and Design Considerations for Multi-Service, Multi-Layer, Multi-Domain Hybrid Networks," *IEEE INFOCOM 2007 HSN Workshop*, Anchorage, Alaska, May 2007.
6. A. Farrell, "A Framework for Inter-Domain Traffic Engineering," *IETF Re-quest for Comments (RFC) 4726*, November 2006.
7. A. Farrell, J. Vasseur, J. Ash, "A Path Computation Element (PCE)-Based Architecture," *IETF Request for Comments (RFC) 4655*, August 2006.
8. I. Iliadis, "Optimal PNNI Complex Node Representations for Restrictive Costs and Minimal Path Computation Time," *IEEE/ACM Transactions on Network-ing*, Vol. 8, No. 4, August 2000.

9. F. Hao, E. Zegura, "On Scalable QoS Routing: Performance Evaluation of Topology Aggregation," *IEEE INFOCOM 2000*, pp. 147–156, 2000.
10. B. Awerbuch, Y. Shavitt, "Topology Aggregation for Directed Graphs," *IEEE/ACM Trans. on Networking*, Vol. 9, No. 2, pp. 82–90, February 2001.
11. K. Liu, K. Nahrstedt, S. Chen, "Routing with Topology Abstraction in Delay-Bandwidth Sensitive Networks," *IEEE/ACM Transactions on Networking*, Vol. 12, No. 1, pp. 17–29, February 2004.
12. X. Yang, B. Ramamurthy, "Inter-Domain Dynamic Routing in Multi-Layer Optical Transport Networks," *IEEE GLOBECOM 2003*, San Francisco, CA, December 2003.
13. Y. Zhu, A. Jukan, M. Ammar, "Multi-Segment Wavelength Routing in Large-Scale Optical Networks," *IEEE ICC 2003*, Anchorage, AL, May 2003.
14. G. Liu, C. Ji, V. W. S. Chan, "On the Scalability of Network Management In-formation for Inter-Domain Light-Path Assessment," *IEEE/ACM Transactions on Networking*, Vol. 13, No. 1, pp. 160–172, January 2005.
15. S. Sanchez-Lopez, et al., "A Hierarchical Routing Approach for GMPLS-Based Control Plane for ASON," *IEEE ICC 2005*, Seoul, Korea, June 2005.
16. Q. Liu, et al., "Hierarchical Inter-Domain Routing and Lightpath Provisioning in Optical Networks," *OSA Journal of Optical Networking*, Vol. 5, No. 10, pp. 764–774, October 2006.
17. P. Ho, H. T. Mouftah, "A Novel Survivable Routing Algorithm for Segment Shared Protection in Mesh WDM Networks with Partial Wavelength Conversion," *IEEE JSAC*, Vol. 22, No. 8, pp. 1539–1548, October 2004.
18. N. Ghani, et al., "Layer 1 VPN Services in Distributed Next-Generation SONET/SDH Networks With Inverse Multiplexing," *OSA Journal of Optical Networking*, Vol. 5, No. 5, pp. 367–382, May 2006.
19. Y. Zhu, et al., "End-to-End Service Provisioning in Multi-Granularity Multi-Domain Optical Networks," *IEEE ICC 2004*, Paris, France, June 2004.
20. R. Alnuweri, et al., "Performance of New Link State Advertisement Mechanisms in Routing Protocols with Traffic Engineering Extensions," *IEEE Comm. Magazine*, Vol. 42, No. 5, pp. 151–162, May 2004.
21. A. Hadjiantonis, et al., "Evolution to a Converged Layer 1, 2 in a Global-Scale, Native Ethernet Over WDM-Based Optical Networking Architecture," to appear in *IEEE Journal of Selected Areas in Communications*, 2007.
22. I. Widjaja, et al., "A New Approach for Automatic Grooming of SONET Circuits To Optical Express Links," *IEEE International Conference on Communications (ICC) 2003*, Anchorage, AL, May 2003.
23. C. Ou, et al., "Traffic Grooming for Survivable WDM Networks-Shared Protection," *IEEE JSAC*, Vol. 21, No. 9, pp. 1367–1383, November 2003.
24. C. Lee, et al., "Exclude Routes-Extensions to RSVP-TE," *IEEE Communications Magazine*, November 2006.

Grooming of Scheduled Demands in Multi-Layer Optical Networks

Maurice Gagnaire, Josué Kuri*, Elias A. Doumith

16.1 Introduction

Introducing multiple switching granularities in transport networks aims in part at reducing their cost by taking advantage of traffic aggregation. As a matter of fact, an appropriate aggregation of low-rate connections into high-rate connections¹ and the use for the latter of a switching technology whose cost-per-bit is lower than for the former can result in lower infrastructure costs. The economic gain, when compared to single switching granularity networks, depends on both traffic-related factors such as the time/space distribution of connections and equipment or network architecture factors such as the cost of ports and matrices for each switching granularity, the topology of the network, the provisioning algorithms used for routing and grooming of connections, etc.

In other chapters of this book, traffic grooming has been investigated in contexts such as protection/restoration, multicast, and multi-domain networks. In this chapter, we consider the problem of efficiently grooming (a) electrical “subwavelength” demands into lightpaths, and (b) lightpaths into wavebands. Whereas most investigations deal with either permanent or random traffic demands, we focus in this chapter on a particular class of traffic demands referred to as scheduled demands, or SxDs, introduced in [1, 2]. A scheduled traffic demand corresponds to a connection or set of connections for which the network operator knows in advance the set-up and tear-down dates. These connections could be requested, for example, by a company to connect its headquarters to its production centers during office hours, or to interconnect its data centers during the night, when database backups are performed. Since companies tend to have well-established operations and procedures, they can provide relatively accurate information to the network operator

* This work was carried out while the author was at ENST Paris. The opinions expressed in this paper are those of the author and not necessarily those of Orange Labs.

¹ For example, Low Order (LO) and High Order (HO) SONET/SDH connections.

about when the bandwidth is really needed. An operator can leverage the knowledge about the timing of demands to make a more effective use of its network resources.² For example, a same set of switch ports and transponders can be allocated to multiple connections routed over a same path if the operator knows in advance that the connections are time-disjoint. Conceptually, a permanent connection is a special case of a scheduled demand for which the set-up and tear-down dates are $-\infty$ and ∞ , respectively. In practice, the set-up and tear-down dates of a permanent connection correspond to the beginning and end dates of the planning period considered in a network design exercise. Therefore, by solving the problem of provisioning scheduled demands, we also solve the problem of provisioning permanent demands.

From a mathematical modeling point of view, an appealing characteristic of scheduled demands is that they capture the dynamic changes of the traffic load in a *deterministic* manner, which eases the use of well-known optimisation techniques such as heuristics [3, 4, 5], meta-heuristics [6], linear programming [7, 8, 9], etc.

A scheduled demand is formally characterized by a tuple (s, d, n, α, β) in which s and d represent the source and destination nodes, n is the requested capacity, and α and β represent the set-up and tear-down dates. Two forms of scheduled demands are considered in this chapter: Scheduled Lightpath Demands (SLDs) and Scheduled Electrical Demands (SEDs), also known as “subwavelength” demands. For SLDs, the capacity n is expressed in number of lightpaths (each one with a nominal rate of either 2.5 or 10 Gbps), and for SEDs, n is expressed as a fraction of a lightpath’s capacity. For example, $n = 0.4$ corresponds to 1 Gbps with 2.5 Gbps lightpaths.

For the problem of grooming SEDs into SLDs, we develop in this chapter a mathematical formulation to quantify the number of Electrical Cross-Connect (EXC) and Wavelength Cross-Connect (WXC) ports required to provision a set of SEDs in a given topology [10, 11, 12, 13]. For the problem of grooming SLDs into wavebands, we develop a different formulation which quantifies the monetary cost of provisioning a set of SLDs in a network in which an integrated Wavelength Cross-Connect (WXC) / Waveband Cross-Connect (BXC) equipment is present in each node.³ We refer to the former as the electrical grooming problem and to the latter as the optical grooming problem.

² This is analogous to what airlines do: passengers that buy tickets well in advance are rewarded with lower fares because early bookings reduce the uncertainty of the future demand, which allows the airline to make a more effective use of its resources.

³ An EXC is essentially the same equipment that has been referred to as a DXC in earlier chapters, and WXC and BXC are both OXCs, operating at different wavelength selectivities. - Editors

16.2 Architecture of Network Nodes

Figure 16.1 depicts the architecture of a multi-granularity switching cross-connect (MG-OXC) that integrates an electrical cross-connect (EXC), a wavelength cross-connect (WXC), and a waveband cross-connect (BXC). The composite signal received on the input fibers of the BXC is demultiplexed to extract individual wavelengths which are in turn multiplexed into wavebands that arrive at the input ports of the BXC switching fabric. The BXC cross-connects input wavebands to either output or drop ports. Dropped wavebands are demultiplexed to extract individual wavelengths which are cross-connected by the WXC to either output or drop ports. Dropped wavelength connections are terminated by the EXC, which extracts “subwavelength” connections (or flows) that are sent to client equipment (e.g., routers) through the EXC drop ports. The MG-OXC has the capability of dropping full lightpaths directly from the WXC to client equipment. In the opposite direction, “subwavelength” connections (or flows) received from client equipment through the EXC add ports are groomed into lightpaths, which are received by the WXC through its add ports and groomed into wavebands that are received by the BXC through its add ports and groomed into fibers. The WXC of the MG-OXC has the capability of adding lightpaths directly from client equipment.

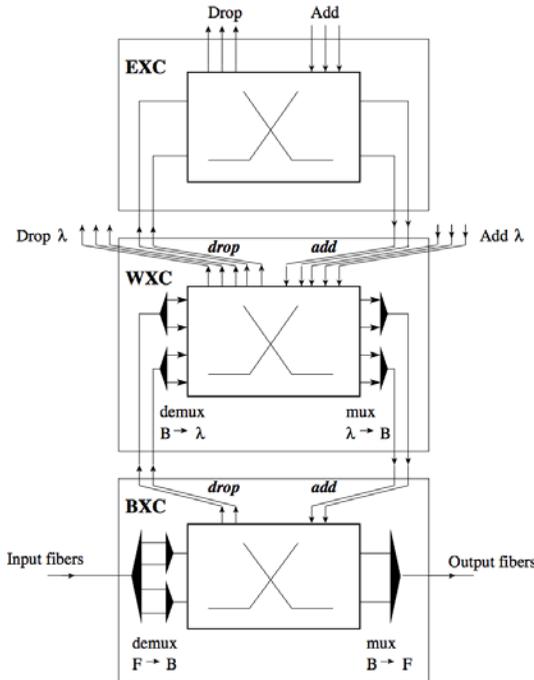


Fig. 16.1. Multilayer node architecture

In this chapter, we deal with electrical and optical grooming as two separate problems. In the former case, the network consists of EXC/WXC nodes (no BXC), whereas in the latter, the network consists of WXC/BXC nodes (no EXCs).

16.3 Rationale of Grooming Strategies

The goal of grooming strategies in multi-layer networks is to reduce the overall management burden and network cost (compared to single-layer networks) by efficiently grouping a large number of low-rate connections into a relatively small number of high-rate connections. The overall network cost can be reduced if the cost-per-bit is lower in the network layer bearing the high-rate connections than in the upper layers, and if the cost of grouping/ungrouping connections does not outweigh the savings of the high-rate connections' layer. Architectures like the one of Fig. 16.1 may reduce the network cost if BXC are less expensive than WXC, and WXC are in turn less expensive than EXC. This tends to be the case with state-of-the-art technology since EXCs require electronic switching fabrics and opto-electronic converters on I/O ports, whereas WXC and BXC require potentially less opto-electronic devices and can be based on less expensive optical switching fabrics.

The presence of transit traffic in a multi-granularity switching cross-connect can reduce the number of required EXC and WXC ports since wavebands bearing transit lightpaths are switched directly from input to output ports of the BXC without requiring any port on the WXC. Likewise, lightpaths bearing transit "sub-wavelength" traffic are switched directly from input to output ports of the WXC without requiring any port on the EXC. Lightpaths on the input ports of the WXC are either recombined into wavebands through the output ports or sent to either client equipment or input ports of the EXC. The EXC extracts "sub-wavelength" connections (or flows) which are sent to client equipment through the drop ports or recombined into new lightpaths sent to the WXC through the output ports.

Provisioning in a multi-layer network involves essentially the definition of both, a path for each SEDs or SLDs, and the way the demands are grouped and ungrouped (potentially several times) along the path. In a single-layer network, a simple strategy consisting in routing connections along the shortest path in terms of hops results in the lowest cost since the total number of required cross-connect ports is minimized. In a multi-layer network, the provisioning aimed at minimizing cost is less obvious since demands can be deviated from their shortest paths in order to create long, efficiently utilized high-rate connections to reduce the overall network cost. The provisioning process becomes even more complex when the demands are SxDs, instead of permanent connections, since the time-disjointness among demands needs to be taken into account in order to reuse network resources as much as possible.

16.3.1 Electrical Grooming

The characteristic of the SEDs lies in the fact that their flow is smaller than the capacity of a lightpath. This particularity is taken into account by use of electrical aggregation. The grooming of multiple SEDs consists in multiplexing their electrical flows onto the same Grooming Lightpath (GL) in an EXC. Grooming requires all the considered SEDs to share at least one common fiber link and all of them to be active during a common period of time. In order to explain the rationale of electrical traffic grooming, let us consider the simple case of the aggregation of two SEDs into a GL. These two SEDs, referred to as δ_1 and δ_2 , are characterized by the tuples $(s_1, d_1, n_1, \alpha_1, \beta_1)$ and $(s_2, d_2, n_2, \alpha_2, \beta_2)$, respectively.

Figure 16.2 illustrates the network topology that supports δ_1 and δ_2 . Below the illustration of the network we have plotted the time diagram of δ_1 and δ_2 . Both requests are active from α_2 to β_1 and they have a common path between node g_1 and node g_2 .

When the network does not have grooming capabilities, each SED is routed using a separate GL between its source and its destination nodes, as illustrated in Fig. 16.3. In other words, δ_1 is routed using the GL φ_1 between s_1 and d_1 , while δ_2 is routed using the GL φ_2 between s_2 and d_2 . As a result, two optical channels are needed on each fiber link between nodes g_1 and g_2 .

On the other hand, when the network has grooming capabilities, δ_1 and δ_2 can be groomed together between nodes g_1 and g_2 if $n_1 + n_2 \leq C_\omega$ as illustrated in Fig. 16.4 (C_ω is the nominal capacity of a lightpath). This grouping creates an aggregated demand $\delta_a = (g_1, g_2, n_1 + n_2, \alpha_2, \beta_1)$ borne by GL φ_a . However, since δ_1 is also active during period $[\alpha_1, \alpha_2]$ and δ_2 is also active during period $[\beta_1, \beta_2]$, two additional requests must be considered: $\delta_f = (g_1, g_2, n_1, \alpha_1, \alpha_2)$ to take into account the remaining active period of δ_1 and $\delta_g = (g_1, g_2, n_2, \beta_1, \beta_2)$

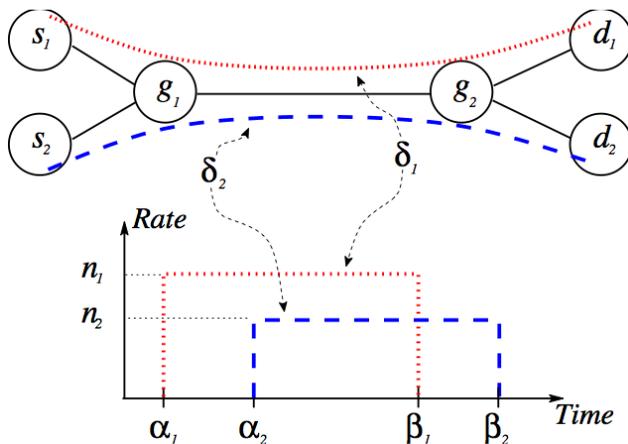


Fig. 16.2. Time diagram of two SEDs

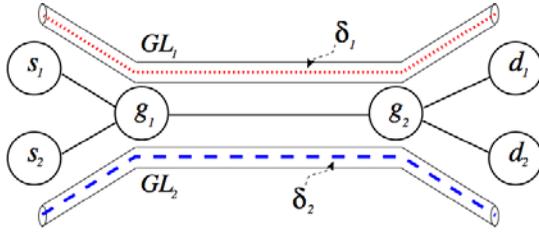


Fig. 16.3. Network without grooming capabilities

to take into account the remaining active period of δ_2 . In addition, we still need to take into account the non-common segments of the paths of δ_1 and δ_2 . This can be done by introducing four additional traffic requests δ_b , δ_c , δ_d , and δ_e :

- Request $\delta_b = (s_1, g_1, n_1, \alpha_1, \beta_1)$ carried by GL \wp_b represents a segment of the path of δ_1 totally disjoint from the path of δ_2 . It connects s_1 , the source node of δ_1 to the grooming node g_1 .
- Request $\delta_c = (g_2, d_1, n_1, \alpha_1, \beta_1)$ carried by GL \wp_c represents another segment of the path of δ_1 totally disjoint from the path of δ_2 . It connects the grooming node g_2 to d_1 , the destination node of δ_1 .

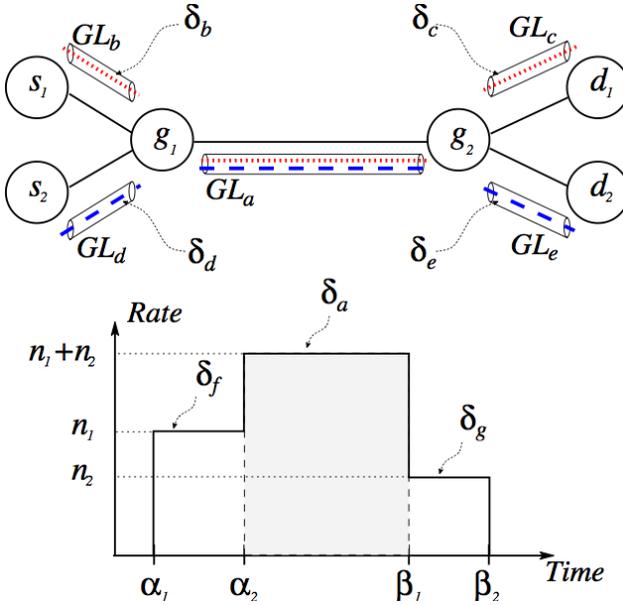


Fig. 16.4. Network with grooming capabilities

- Request $\delta_d = (s_2, g_1, n_2, \alpha_2, \beta_2)$ carried by GL φ_d represents a segment of the path of δ_2 totally disjoint from the path of δ_1 . It connects s_2 , the source node of δ_2 , to the grooming node g_1 .
- Request $\delta_e = (g_2, d_2, n_2, \alpha_2, \beta_2)$ represents another segment of the path of δ_2 totally disjoint from the path of δ_1 . It connects the grooming node g_2 to d_2 , the destination node of δ_2 .

To sum up, grooming together requests δ_1 and δ_2 creates the aggregated request δ_a and a set of six additional requests. Some of these requests ($\delta_b, \delta_c, \delta_d$, and δ_e) take into account the non-common segments of the paths of δ_1 and δ_2 . The remaining requests (δ_f and δ_g) take into account the non-common periods of time. Depending on the characteristics of the initial requests δ_1 and δ_2 , some of these additional requests may not exist. However, those which do exist form the set of marginal demands. As a result of the grooming process, when grooming two SEDs, the set of all the SEDs must be modified by adding the aggregated request δ_a and the set of marginal demands and by removing the original requests δ_1 and δ_2 .

At node g_1 , the two demands δ_1 and δ_2 must be switched from the WXC to the EXC in order to be groomed together. Hence two receiving r_3 electrical ports and two emitting o_3 ports are needed at this node. Once groomed together onto the the GL φ_a , the resulting grooming lightpath is switched back to the WXC using an emitting e_3 electrical port and a receiving o_3 optical port. Arriving at node g_2 , the GL φ_a is switched from the WXC to the EXC using an emitting o_3 optical port and a receiving r_3 electrical port. At the EXC of node g_2 , the initial requests δ_1 and δ_2 are rebuilt and each one is switched to its destination using two emitting e_3 electrical ports and two receiving o_3 optical ports. In brief, in order to groom these two requests, six additional electrical ports and six additional optical ports are needed. Figure 16.5 details the ports involved in the grouping and ungrouping operations at nodes g_1 and g_2 , respectively.

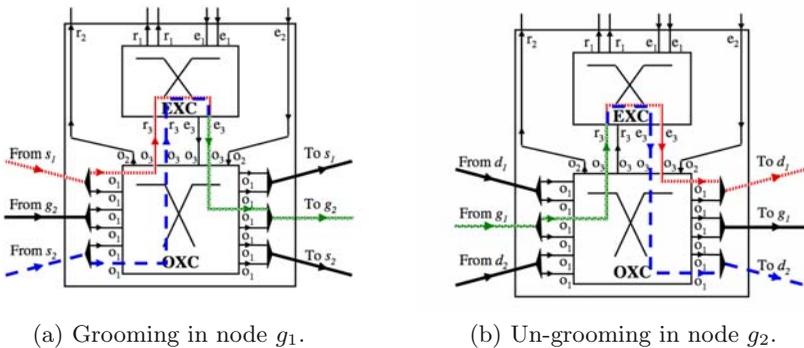


Fig. 16.5. Grooming operations in nodes g_1 and g_2

16.3.2 Optical Grooming

For the optical grooming problem, we consider a network of WXC/BXC nodes (no EXCs). We assume that WXCs provide full wavelength conversion. This simplifies wavelength assignment since the wavelength continuity constraint does not exist. As explained in Section 16.2, the lightpaths at the output (input) ports of the WXC are multiplexed (demultiplexed) into (from) wavebands which are directly added in (dropped from) a BXC. Thus, the WXCs are not directly connected to each other but through waveband-switching connections between BXCs. This means that lightpaths are instantiated over a logical topology of waveband-switching connections. These connections cannot be directly added in or dropped from the BXC. The number of WXC I/O ports is a multiple of the waveband size.

Figure 16.6 shows how three lightpaths can be instantiated over a logical topology formed by two waveband-switching connections, one between nodes 1 and 3 and another between nodes 3 and 4. Two lightpaths are added in the WXC of node 1 using the WXC add ports. They are then multiplexed into a waveband-switching connection, which is added into the BXC of the same node using one BXC add port. The resulting waveband-switching connection is multiplexed into a fiber. In node 2, the connection is demultiplexed from the incoming fiber, switched to an output port and multiplexed into a fiber. In node 3, the connection is demultiplexed from the incoming fiber and dropped using one BXC drop port. This connection is demultiplexed in order to retrieve the two lightpaths which are then switched from the input to the output ports of the node. The lightpaths are multiplexed together with a third lightpath added in node 3 into a waveband-switching connection which is added in the BXC of the node. The connection is multiplexed into a fiber. In node 4, the connection is demultiplexed from the incoming fiber and dropped. The three

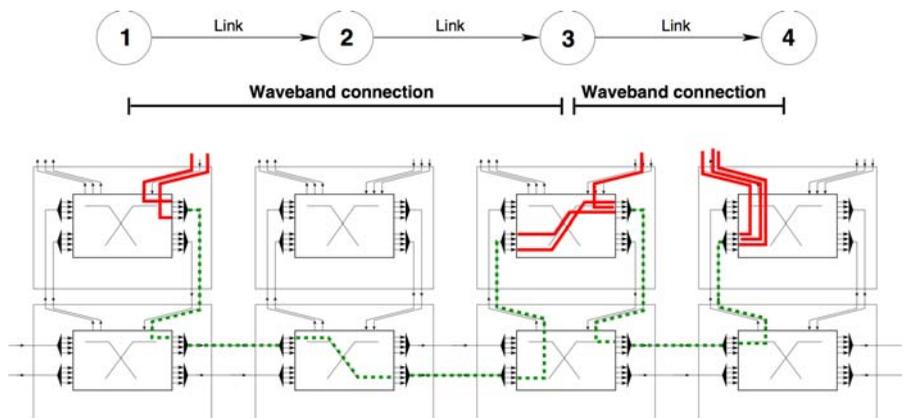


Fig. 16.6. A possible configuration of MG-OXCs used to set up three lightpaths

lightpaths are demultiplexed from the waveband-switching connection and dropped using the WXC drop ports of the node.

Waveband-switching introduces an intermediate network layer between the physical network and the SLDs. A logical topology of waveband-switching connections must be defined and mapped on the physical network. An SLD is instantiated by defining a path on the logical topology and assigning waveband-switching connections on each arc of the path to the SLD's lightpaths. In this context, grooming refers to the aggregation (and disaggregation) of lightpaths into waveband-switching connections of the logical topology. Thus, the optical grooming problem under consideration consists in defining a logical topology of waveband-switching connections, routing these connections on a physical network, routing a set of SLDs on the logical topology, and assigning the waveband-switching connections of the logical topology to the SLDs such that the network cost is minimal. The cost of the network is equal to the sum of the costs of the MG-OXC required to implement the solution. The problem is schematically illustrated by Fig. 16.7. An instance of the problem is defined by a set of SLDs and the topology of a physical network. The MG-OXC cost function models the cost of a MG-OXC as a function of the number of WXC and BXC ports of the MG-OXC (the function is defined in Equation 16.16). The network functional model defines the architecture of the switches, including the definition of the size of the wavebands (the number of lightpaths that can be groomed into a waveband-switching connection). The objective function defines the optimality criterion to be satisfied, for example, the minimization of the network cost. The solution to the problem consists of a set of Routed Scheduled Band Groups (RSBGs) and of an assignment of SLDs to RSBGs. An RSBG is similar to an SLD in that it is defined by a tuple $(s, d, n, \alpha, \beta, P)$ where s and d are the source and destination BXC of the RSBG, n is the number of waveband-switching connections in the group, α and β are the set-up and tear-down dates of the RSBG, and P is the route in the physical network. The set-up date α (tear-down date β) of an RSBG is the earliest set-up (latest tear-down) date of any of the SLDs assigned to this RSBG.

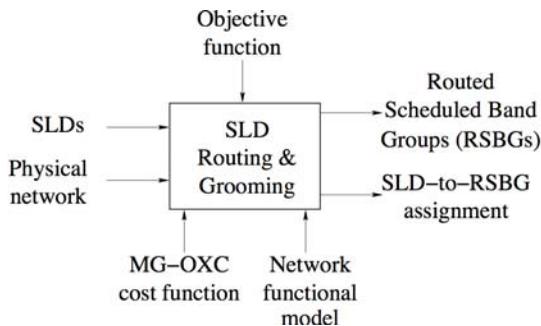


Fig. 16.7. Schematic representation of the optical grooming problem

16.4 Mathematical Formulation

In this section, we develop the mathematical formulations used to quantify the network cost in the electrical and optical grooming problems. The notations in Sections 16.4.1 and 16.4.2 are independent, that is, the symbols used in one subsection are not related to the symbols of the other.

16.4.1 Electrical Grooming

The design of a grooming-capable network is formally stated below. The inputs to the problem are:

- A physical topology represented by an arc-weighted symmetrical directed graph $G = (V; E; w)$. $V = \{v_1, v_2, \dots, v_N\}$ is the set of network nodes and $E = \{e_1, e_2, \dots, e_F\}$ is the set of physical links interconnecting the nodes. Nodes correspond to EXC/WXC multi-layer switches and links correspond to the fibers of the network. Though links are unidirectional, we assume that there are equal number of fibers linking two nodes in different directions. Links are assigned weights $w : E \mapsto \mathbb{R}^+$ which may correspond to the physical length or cost of the links. $N = |V|$ and $F = |E|$ represent the number of nodes and links in the network, respectively.
- A set $\mathcal{D} = \{\delta_1, \delta_2, \dots, \delta_M\}$ of M SxD requests to be routed on the network. These requests can be divided into a set $\mathcal{D}^L = \{\delta_1^L, \delta_2^L, \dots, \delta_{M^L}^L\}$ of M^L SLD requests and a set $\mathcal{D}^E = \{\delta_1^E, \delta_2^E, \dots, \delta_{M^E}^E\}$ of M^E SED requests. SLDs (δ_i^L) and SEDs (δ_i^E) are represented by a tuple $(s_i, d_i, n_i, \alpha_i, \beta_i)$ where $s_i, d_i \in V$ are the source and destination nodes of the demand, n_i is the traffic rate requested by the connection, and α_i and β_i are its setup and tear-down dates. For SLDs, n_i is an integer, while for SEDs n_i is a real number smaller than 1.
- A set of available routes $\mathcal{P}_i = \{P_{i,1}, P_{i,2}, \dots, P_{i,K}\}$ connecting the source node s_i to the destination node d_i of request δ_i^L or δ_i^E . For each request, we compute beforehand K alternate shortest paths in terms of effective length connecting its source node to its destination node using the algorithm described in [14] (if as many paths exist, otherwise we consider the available ones).

The set-up and tear-down dates of all these requests are to be taken into account because they represent time instants when a change in the traffic flow is observed. Let \mathcal{E} be the ordered set grouping the set-up and tear-down dates of all the requests in the set \mathcal{D} . Because some requests may have the same set-up and tear-down dates, the number $\mathbb{E} = |\mathcal{E}|$ of time instants in \mathcal{E} is less than or equal to $2 \times M$.

$$\begin{aligned} \mathcal{E} &= \left(\bigcup_{i=1}^M \alpha_i \right) \cup \left(\bigcup_{i=1}^M \beta_i \right) \\ &= \{\epsilon_1, \epsilon_2, \dots, \epsilon_{\mathbb{E}}\} \text{ such as } \epsilon_1 < \epsilon_2 < \dots < \epsilon_{\mathbb{E}} \end{aligned} \quad (16.1)$$

In the case of SEDs, a physical path $P_{i,\rho_i} \in \mathcal{P}_i$ is assigned in a first step to the request δ_i^E ($1 \leq i \leq M^E$). P_{i,ρ_i} is the ρ_i^{th} shortest path between s_i and d_i ($1 \leq \rho_i \leq K$). Once the physical path P_{i,ρ_i} is chosen, we determine in a second step the intermediate nodes of P_{i,ρ_i} in which δ_i^E must be switched to the electrical level for grooming. In this way, we obtain the set of grooming lightpaths that the request traverses in the logical topology.

A physical path P_{i,ρ_i} , composed of ℓ_{i,ρ_i} hops, has $(\ell_{i,\rho_i} - 1)$ intermediate nodes and can be decomposed into $\mathbb{L}_{i,\rho_i} = 2^{\ell_{i,\rho_i} - 1}$ different sets of grooming lightpaths. For example, the physical path $P_{i,\rho_i} = 1 - 2 - 3 - 4$ between nodes 1 and 4 can be decomposed into the following $2^{3-1} = 4$ different sets of grooming lightpaths:

- $L_{i,\rho_i,1} = \{[1 - 2 - 3 - 4]\}$
- $L_{i,\rho_i,2} = \{[1 - 2], [2 - 3 - 4]\}$
- $L_{i,\rho_i,3} = \{[1 - 2 - 3], [3 - 4]\}$
- $L_{i,\rho_i,4} = \{[1 - 2], [2 - 3], [3 - 4]\}$

$L_{i,\rho_i,\lambda_i,\rho_i}$ is the λ_i^{th} possible decomposition of the physical path P_{i,ρ_i} into a set of grooming lightpaths ($1 \leq \lambda_i,\rho_i \leq \mathbb{L}_{i,\rho_i}$). Let $\mathcal{L}_{i,\rho_i} = \{L_{i,\rho_i,1}, L_{i,\rho_i,2}, \dots, L_{i,\rho_i,\mathbb{L}_{i,\rho_i}}\}$ be the set of all possible decompositions of the path P_{i,ρ_i} into its inherent sets of grooming lightpaths.

The solution of the optimization problem consists in assigning to the request δ_i^E a physical path $P_{i,\rho_i} \in \mathcal{P}_i$ and a set of grooming lightpaths $L_{i,\rho_i,\lambda_i,\rho_i} \in \mathcal{L}_{i,\rho_i}$. This solution is represented by $\Pi_{\mathcal{D}^E,\rho,\lambda} = \{(P_{1,\rho_1}, L_{1,\rho_1,\lambda_{1,\rho_1}}), (P_{2,\rho_2}, L_{2,\rho_2,\lambda_{2,\rho_2}}), \dots, (P_{M^E,\rho_{M^E}}, L_{M^E,\rho_{M^E},\lambda_{M^E,\rho_{M^E}}})\}$.

Our objective is to satisfy all the requests at the lowest cost. The network cost is expressed in terms of the number of electrical ports and of optical ports used and can be obtained by means of mathematical operations based on matrix algebra.

Computation of Network Cost

For an instance, $\Pi_{\mathcal{D}^E,\rho,\lambda} = \{(P_{1,\rho_1}, L_{1,\rho_1,\lambda_{1,\rho_1}}), (P_{2,\rho_2}, L_{2,\rho_2,\lambda_{2,\rho_2}}), \dots, (P_{M^E,\rho_{M^E}}, L_{M^E,\rho_{M^E},\lambda_{M^E,\rho_{M^E}}})\}$ of the network design problem under SED traffic requests, we define the following matrices:

- **Request Matrix:** The request matrix, denoted by $\underline{\Delta}^E = \{\delta_{i,t}^E\}$, represents the SED traffic requests over time. $\underline{\Delta}^E$ is a $[M^E \times \mathbb{E}]$ matrix. An element $\delta_{i,t}^E$ of such a matrix is a binary value ($\delta_{i,t}^E \in \{0, 1\}$) specifying the presence ($\delta_{i,t}^E = 1$) or absence ($\delta_{i,t}^E = 0$) of the SED request $\delta_i^E(s_i, d_i, \alpha_i, \beta_i, n_i)$ at time instant ϵ_t .

$$\delta_{i,t}^E = \begin{cases} 1 & \text{if } \alpha_i \leq \epsilon_t < \beta_i, \\ 0 & \text{otherwise.} \end{cases} \tag{16.2}$$

- **Source/Destination Matrix:** The source matrix, denoted by $\underline{S}^E = \{s_{i,n}^E\}$, represents the source nodes of the SED requests. The destination matrix, denoted by $\underline{D}^E = \{d_{i,n}^E\}$, represents the destination nodes of the SED requests. \underline{S}^E and \underline{D}^E are $[M^E \times N]$ matrices. An element $s_{i,n}^E$ ($d_{i,n}^E$) of such a matrix is a binary value specifying if the SED request $\delta_i^E(s_i, d_i, \alpha_i, \beta_i, n_i)$ has node v_n as source (destination) node or not.

$$s_{i,n}^E = \begin{cases} 1 & \text{if } s_i = v_n, \\ 0 & \text{otherwise.} \end{cases} \quad d_{i,n}^E = \begin{cases} 1 & \text{if } d_i = v_n, \\ 0 & \text{otherwise.} \end{cases} \quad (16.3)$$

As shown in Section 16.3.1, when grooming two SED requests $\delta_{i_1}^E$ and $\delta_{i_2}^E$, the set \mathcal{D}^E of all the SEDs must be modified by adding the aggregated request and the set of marginal demands, and by removing the original requests $\delta_{i_1}^E$ and $\delta_{i_2}^E$. Consequently, the number of SED requests as well as their characteristics are modified. Let $\tilde{\mathcal{D}}^E = \{\tilde{\delta}_1^E, \tilde{\delta}_2^E, \dots, \tilde{\delta}_{M^E}^E\}$ be the set of the \tilde{M}^E groomed SED requests obtained at the end of the grooming optimization process. Each groomed SED request $\tilde{\delta}_i^E \in \tilde{\mathcal{D}}^E$ is routed over a unique path \tilde{P}_i^E . As each element $\tilde{\delta}_i^E$ can be either a single low-speed connection δ_j^E or the concatenation of two or several low-speed connections $\tilde{\delta}_i^E = \uplus_j \delta_j^E$, its corresponding path is determined according to the grooming lightpath sets of its lower-speed connections ($\tilde{P}_i^E \in L_{j,\rho_j,\lambda_j,\rho_j}$).

Based on the new SED set $\tilde{\mathcal{D}}^E$, we define the following additional matrices:

- **Groomed Request Matrix:** The groomed request matrix, denoted by $\underline{\tilde{\Delta}}^E = \{\tilde{\delta}_{i,t}^E\}$, represents the groomed SED traffic requests over time. $\underline{\tilde{\Delta}}^E$ is a $[\tilde{M}^E \times \mathbb{E}]$ matrix. An element $\tilde{\delta}_{i,t}^E$ of such a matrix is an integer number due to the fact that the groomed SED demand $\tilde{\delta}_i^E(\tilde{s}_i, \tilde{d}_i, \tilde{\alpha}_i, \tilde{\beta}_i, \tilde{n}_i)$ can be the concatenation of several requests of smaller rates. Consequently, it could require more than one lightpath to carry its traffic load. $\tilde{\delta}_{i,t}^E$ is equal to the smallest integer greater than or equal to the traffic required by the groomed SED request $\tilde{\delta}_i^E$ and specifies the number of lightpaths used by this request at time instant ϵ_t .

$$\tilde{\delta}_{i,t}^E = \begin{cases} \lceil \tilde{n}_i \rceil & \text{if } \tilde{\alpha}_i \leq \epsilon_t < \tilde{\beta}_i, \\ 0 & \text{otherwise.} \end{cases} \quad (16.4)$$

- **Groomed Routing Matrix:** The groomed routing matrix, denoted by $\underline{\tilde{R}}^E = \{\tilde{r}_{i,f}^E\}$, represents the use of the physical links by the groomed SED requests. $\underline{\tilde{R}}^E$ is a $[\tilde{M}^E \times F]$ matrix. An element $\tilde{r}_{i,f}^E$ of such a matrix is a binary value ($\tilde{r}_{i,f}^E \in \{0, 1\}$) specifying if the physical path \tilde{P}_i^E assigned to the groomed SED request $\tilde{\delta}_i^E(\tilde{s}_i, \tilde{d}_i, \tilde{\alpha}_i, \tilde{\beta}_i, \tilde{n}_i)$ passes through the physical link e_f ($\tilde{r}_{i,f}^E = 1$) or not ($\tilde{r}_{i,f}^E = 0$).

$$\tilde{r}_{i,f}^E = \begin{cases} 1 & \text{if } e_f \in \tilde{P}_i^E, \\ 0 & \text{otherwise.} \end{cases} \quad (16.5)$$

- **Groomed Source/Destination Matrix:** The groomed source matrix, denoted by $\tilde{\underline{S}}^E = \{\tilde{s}_{i,n}^E\}$, represents the source nodes of the groomed SED requests. The groomed destination matrix, denoted by $\tilde{\underline{D}}^E = \{\tilde{d}_{i,n}^E\}$, represents the destination nodes of the groomed SED requests. $\tilde{\underline{S}}^E$ and $\tilde{\underline{D}}^E$ are $[\tilde{M}^E \times N]$ matrices. An element $\tilde{s}_{i,n}^E$ ($\tilde{d}_{i,n}^E$) of such a matrix is a binary value specifying if the groomed SED request $\tilde{\delta}_i^E(\tilde{s}_i, \tilde{d}_i, \tilde{\alpha}_i, \tilde{\beta}_i, \tilde{n}_i)$ has node v_n as source (destination) node or not.

$$\tilde{s}_{i,n}^E = \begin{cases} 1 & \text{if } \tilde{s}_i = v_n, \\ 0 & \text{otherwise.} \end{cases} \quad \tilde{d}_{i,n}^E = \begin{cases} 1 & \text{if } \tilde{d}_i = v_n, \\ 0 & \text{otherwise.} \end{cases} \quad (16.6)$$

By multiplying the transposed request matrix $\underline{\Delta}^{E^T}$ by the source matrix \underline{S}^E a new matrix $\underline{\Psi}^{E,Em} = \{\psi_{t,n}^{E,Em}\}$ is obtained. $\underline{\Psi}^{E,Em}$, of dimension $[\mathbb{E} \times N]$, represents the use of sub-wavelength ‘add’ ports at the nodes. An element $\psi_{t,n}^{E,Em}$ of such a matrix is an integer value representing the number of emitting e_1 electrical ports in use at node v_n at each instant ϵ_t . For a given node v_n , let $v_n^{E,Em}$ be the maximum value of the nodes’ add ports $\psi_{t,n}^{E,Em}$ evaluated over the whole observation period. $v_n^{E,Em}$ gives the number of emitting e_1 electrical ports installed at the node v_n . The number $\gamma_{e_1}^E$ of emitting e_1 electrical ports in the network is the sum of $v_n^{E,Em}$ over all the network nodes.

$$\underline{\Psi}^{E,Em} = \underline{\Delta}^{E^T} \times \underline{S}^E \quad (16.7a)$$

$$v_n^{E,Em} = \max_{1 \leq t \leq \mathbb{E}} \psi_{t,n}^{E,Em} \quad (16.7b)$$

$$\gamma_{e_1}^E = \sum_{n=1}^N v_n^{E,Em} \quad (16.7c)$$

Similarly, the number $\gamma_{r_1}^E$ of receiving r_1 electrical ports in the network is computed as follows:

$$\underline{\Psi}^{E,Re} = \underline{\Delta}^{E^T} \times \underline{D}^E \quad (16.8a)$$

$$v_n^{E,Re} = \max_{1 \leq t \leq \mathbb{E}} \psi_{t,n}^{E,Re} \quad (16.8b)$$

$$\gamma_{r_1}^E = \sum_{n=1}^N v_n^{E,Re} \quad (16.8c)$$

By multiplying the transposed groomed request matrix $\tilde{\Delta}^{E^T}$ by the groomed routing matrix \tilde{R}^E , a new matrix $\tilde{\Phi}^E = \{\tilde{\phi}_{t,f}^E\}$ is obtained. $\tilde{\Phi}^E$, of dimension $[\mathbb{E} \times F]$, represents the traffic load over time carried by the links. An element $\tilde{\phi}_{t,f}^E$ of such a matrix is an integer value representing the number of WDM optical channels carried by link e_f at each instant e_t . For a given link e_f , let $\tilde{\varphi}_f^E$ be the maximum value of this traffic $\tilde{\phi}_{t,f}^E$ evaluated over the whole observation period. $\tilde{\varphi}_f^E$ gives the number of WDM optical channels used on this link. As the o_1 optical ports go by pair, one port at each end of a WDM channel, $\tilde{\varphi}_f^E$ represents also the number of o_1 optical port pairs installed at both ends of the link e_f . The number $\gamma_{o_1}^E$ of o_1 optical ports in the network is twice the sum of $\tilde{\varphi}_f^E$ over all the network links. The maximum value of $\tilde{\varphi}_f^E$ ($\forall f \setminus 1 \leq f \leq F$) represents the congestion in the network, i.e., the number of active channels on the most loaded link.

$$\tilde{\Phi}^E = \tilde{\Delta}^{E^T} \times \tilde{R}^E \tag{16.9a}$$

$$\tilde{\varphi}_f^E = \max_{1 \leq t \leq \mathbb{E}} \tilde{\phi}_{t,f}^E \tag{16.9b}$$

$$\gamma_{o_1}^E = 2 \times \sum_{f=1}^F \tilde{\varphi}_f^E \tag{16.9c}$$

By multiplying the transposed groomed request matrix $\tilde{\Delta}^{E^T}$ by the groomed source matrix \tilde{S}^E a new matrix $\tilde{\Psi}^{E,Em} = \{\tilde{\psi}_{t,n}^{E,Em}\}$ is obtained. $\tilde{\Psi}^{E,Em}$, of dimension $[\mathbb{E} \times N]$, represents the use of switching ports from the EXC to the WXC at the nodes. An element $\tilde{\psi}_{t,n}^{E,Em}$ of such a matrix is an integer value representing the number of emitting e_3 electrical ports in use at node v_n at each instant e_t . For a given node v_n , let $\tilde{v}_n^{E,Em}$ be the maximum value of the nodes' switching ports $\tilde{\psi}_{t,n}^{E,Em}$ evaluated over the whole observation period. $\tilde{v}_n^{E,Em}$ gives the number of emitting e_3 electrical ports installed at the node v_n . The number $\gamma_{e_3}^E$ of emitting e_3 electrical ports in the network is the sum of $\tilde{v}_n^{E,Em}$ over all the network nodes.

$$\tilde{\Psi}^{E,Em} = \tilde{\Delta}^{E^T} \times \tilde{S}^E \tag{16.10a}$$

$$\tilde{v}_n^{E,Em} = \max_{1 \leq t \leq \mathbb{E}} \tilde{\psi}_{t,n}^{E,Em} \tag{16.10b}$$

$$\gamma_{e_3}^E = \sum_{n=1}^N \tilde{v}_n^{E,Em} \tag{16.10c}$$

Similarly, the number $\gamma_{r_3}^E$ of receiving r_3 electrical ports in the network is computed as follows:

$$\underline{\tilde{\Psi}}^{E,Re} = \underline{\tilde{\Delta}}^{E^T} \times \underline{\tilde{D}}^E \quad (16.11a)$$

$$\tilde{v}_n^{E,Re} = \max_{1 \leq t \leq \mathbb{E}} \tilde{\psi}_{t,n}^{E,Re} \quad (16.11b)$$

$$\gamma_{r_3}^E = \sum_{n=1}^N \tilde{v}_n^{E,Re} \quad (16.11c)$$

Consequently, the number $\gamma_{o_3}^E$ of o_3 optical ports in the network is equal to:

$$\gamma_{o_3}^E = \gamma_{e_3}^E + \gamma_{r_3}^E \quad (16.12)$$

As a set of SEDs does not have any SLD component to be added or dropped at a node, the number of e_2/r_2 electrical ports and the number of o_2 optical ports are null.

$$\gamma_{e_2}^E = \gamma_{r_2}^E = \gamma_{o_2}^E = 0 \quad (16.13)$$

Finally, the total number ϱ^E of optical ports and the total number ε^E of electrical ports are given by:

$$\begin{aligned} \varrho^E &= \gamma_{o_1}^E + \gamma_{o_2}^E + \gamma_{o_3}^E \\ &= \gamma_{o_1}^E + \gamma_{o_3}^E \end{aligned} \quad (16.14a)$$

$$\begin{aligned} \varepsilon^E &= \gamma_{e_1}^E + \gamma_{e_2}^E + \gamma_{e_3}^E + \gamma_{r_1}^E + \gamma_{r_2}^E + \gamma_{r_3}^E \\ &= \gamma_{e_1}^E + \gamma_{e_3}^E + \gamma_{r_1}^E + \gamma_{r_3}^E \end{aligned} \quad (16.14b)$$

κ being the ratio of the cost of an electrical port to the cost of an optical port, the cost ζ^E of routing the SED requests is then expressed as:

$$\zeta^E = \varrho^E + \kappa \times \varepsilon^E \quad (16.15)$$

16.4.2 Optical Grooming

$G = (V, E, w)$

is an arc-weighted symmetrical directed graph with vertex set $V = \{v_1, v_2, \dots, v_N\}$, arc set $E = \{e_1, e_2, \dots, e_F\}$, and arc weight function $w : E \rightarrow \mathbb{R}_+$. The graph represents a physical telecommunications network. The set V corresponds to the nodes of the network and the set of arcs E to the links interconnecting the nodes. Function w corresponds to the physical length or to the cost of the links (defined, for example, by the network operator).

$N = |V|$, $F = |E|$, K , L , B

are, respectively, the number of nodes and links in the network, the maximum number of possible alternate paths for each demand, the maximum number of possible layouts for each path and the size of a waveband-switching connection

(expressed in number of lightpaths). All the waveband-switching connections in the network have the same size.

$$P = (x_0, x_1, \dots, x_z)$$

describes a *path* in G that is composed of z links $(x_0, x_1), (x_1, x_2), \dots, (x_{z-1}, x_z)$ where the $(x_{i-1}, x_i) \in E$ are all distinct (i.e., the path is loop free).

$$\Delta = \{\delta_1, \delta_2, \dots, \delta_M\}$$

is a set of M Scheduled Lightpath Demands (SLDs), where:

$$\delta_i = (s_i, d_i, n_i, \alpha_i, \beta_i)$$

is a tuple representing SLD number i ; $s_i, d_i \in V$ are the source and destination nodes of the demand, n_i is the number of requested lightpaths, and α_i and β_i are the set-up and tear-down dates.

$$(G, \Delta)$$

is a pair representing an instance of the optical grooming problem.

$$P_{i,k} = (x_0^{i,k}, x_1^{i,k}, \dots, x_{z_i,k}^{i,k}), 1 \leq i \leq M, 1 \leq k \leq K$$

is the k -th alternate path for SLD δ_i from $x_0^{i,k} = s_i$ to $x_{z_i,k}^{i,k} = d_i$. For the purposes of this model, we compute the K physically⁴ shortest paths (if so many exist) for each demand using the algorithm defined in [14]. However, the paths might be defined according to any other criterion (i.e., the function w may map any other value than the length of the links).

$$C : \Pi_\Delta \rightarrow \mathbb{R}_+$$

is the function that computes the cost of an admissible solution $\pi_{\rho,\nu,\Delta}$ to the optical grooming problem (explained below). In order to formalize this function we define the following additional notations.

$$\theta = (\theta_{ij})$$

is a $\{0, 1\}^{M \times M}$ upper triangular matrix; θ_{ij} , $i \leq j$, indicates whether the SLDs δ_i and δ_j overlap in time ($\theta_{ij} = 1$) or not ($\theta_{ij} = 0$). By definition $\theta_{ii} = 1$, $1 \leq i \leq M$, and $\theta_{ij} = 0$ for $i > j$. This matrix expresses the temporal interdependence between the SLDs.

$$\tau = (\tau_{ij}) = \text{diag}(n_i)$$

is a diagonal matrix in which $\tau_{ii} = n_i$, $1 \leq i \leq M$, that is, τ_{ii} is the number of lightpaths required by the SLD δ_i .

$$\mathcal{I} = (\mathcal{I}_{ij})$$

is a $\{0, 1\}^{N \times F}$ matrix; \mathcal{I}_{ij} indicates whether vertex v_i is the termination vertex of arc e_j ($\mathcal{I}_{ij} = 1$) or not ($\mathcal{I}_{ij} = 0$).

⁴ The function w maps the length of each link.

$$\mathcal{O} = (\mathcal{O}_{ij})$$

is a $\{0, 1\}^{N \times F}$ matrix; \mathcal{O}_{ij} indicates whether vertex v_i is the source vertex of arc e_j ($\mathcal{O}_{ij} = 1$) or not ($\mathcal{O}_{ij} = 0$).

$$\mathcal{T} = (t_{ij})$$

is a $\{0, 1\}^{M \times N}$ matrix; t_{ij} indicates whether vertex v_j is the source node of SLD δ_i ($t_{ij} = 1$) or not ($t_{ij} = 0$).

$$\mathcal{U} = (u_{ij})$$

is a $\{0, 1\}^{M \times N}$ matrix; u_{ij} indicates whether vertex v_j is the destination node of SLD δ_i ($u_{ij} = 1$) or not ($u_{ij} = 0$).

$$\mathcal{D} = \theta \cdot \tau \cdot \mathcal{T} = (\mathcal{D}_{ij})_{1 \leq i \leq M, 1 \leq j \leq N},$$

$$\mathcal{D}_j^* = \max_{1 \leq i \leq M} \{\mathcal{D}_{ij}\}$$

\mathcal{D}^* is an N -dimensional vector; \mathcal{D}_j^* is the maximum number of simultaneously active lightpaths (i.e., time-overlapping lightpaths) originating at node v_i .

$$\mathcal{E} = \theta \cdot \tau \cdot \mathcal{U} = (\mathcal{E}_{ij})_{1 \leq i \leq M, 1 \leq j \leq N},$$

$$\mathcal{E}_j^* = \max_{1 \leq i \leq M} \{\mathcal{E}_{ij}\}$$

\mathcal{E}^* is an N -dimensional vector; \mathcal{E}_j^* is the maximum number of simultaneously active lightpaths (i.e., time-overlapping lightpaths) terminating at node v_i .

$$\psi : \mathbb{N} \rightarrow \mathbb{R}_+$$

is the function that determines the cost of a cross-connect (either a WXC or a BXC) as a function of its number of ports. In this work, we consider the function:

$$\psi(x) = a + bx^c \quad a, b \in \mathbb{R}_+, c \in [1, 2[. \quad (16.16)$$

The function captures various technology-specific factors that have an effect on the cost of a cross-connect. The parameter a represents the fixed cost of installing the switch (shelf, power and ventilation systems, etc.). The parameter b represents the cost of a port. Finally, the parameter c accounts for the impact on the cost of the increasing implementation complexity of switching matrices with a large number of ports. This parameter is limited to take values smaller than 2 because, for higher values, it is in general more economical to stack multiple small switches than building a large one when a significant number of ports is required.

$$\lambda = \{\lambda^n\}$$

is a partition of the set of arcs describing path P . We call λ a *layout* and an element λ^n of this partition a *subpath*. For example, a layout of path $(x_0, x_1, x_2, x_3, x_4)$ is $\lambda = \{(x_0, x_1, x_2), (x_2, x_3, x_4)\}$. This layout has subpaths $\lambda^1 = (x_0, x_1, x_2)$ and $\lambda^2 = (x_2, x_3, x_4)$. Another layout is $\{(x_0, x_1), (x_1, x_2, x_3, x_4)\}$. The elements of a subpath must be contiguous arcs of P .

$$\Lambda_{i,k} = \{\lambda_{i,k,j}\}, 1 \leq i \leq M, 1 \leq k \leq K, 1 \leq j \leq L$$

is the set of layouts available for path $P_{i,k}$; $\lambda_{i,k,j}$ is the j -th layout of the path. We assume that there are at most L different layouts defined for a path.

$$\pi_{\rho,\nu,\Delta} = ((P_{1,\rho_1}, \lambda_{1,\rho_1,\nu_1}), (P_{2,\rho_2}, \lambda_{2,\rho_2,\nu_2}), \dots, (P_{M,\rho_M}, \lambda_{M,\rho_M,\nu_M})),$$

$$\rho \in \{1, \dots, K\}^M, \nu \in \{1, \dots, L\}^M$$

is called an *admissible optical grooming solution* for Δ . ρ and ν are M -dimensional vectors whose elements can take values in $[1, K]$ and $[1, L]$, respectively. $\pi_{\rho,\nu,\Delta}$ is fully characterized by the pair (ρ, ν) . An admissible optical grooming solution $\pi_{\rho,\nu,\Delta}$ defines for each SLD, a path $P_{i,k}$ and a layout $\lambda_{i,k,j}$ of this path. A subpath λ^n can be part of several layouts defined in $\pi_{\rho,\nu,\Delta}$. We call the association of λ^n to the subset $\Delta' \subseteq \Delta$ of SLDs whose layouts in $\pi_{\rho,\nu,\Delta}$ share λ^n , a Routed Scheduled Band Group (RSBG). Thus, besides a path and a layout for each SLD, an admissible optical grooming solution $\pi_{\rho,\nu,\Delta}$ also defines a set of RSBGs.

$$\Pi_{\Delta} = \{\pi_{\rho,\nu,\Delta}, \rho \in \{1, \dots, K\}^M, \nu \in \{1, \dots, L\}^M\}$$

is the set of all admissible optical grooming solutions for Δ . Its cardinality is $|\Pi_{\Delta}| = (KL)^M$ (assuming that K paths and L layouts are available for each path; otherwise, $|\Pi_{\Delta}| < (KL)^M$). The set represents the solution space of an optical grooming problem instance (G, Δ) .

$$\mathcal{S}_{\rho,\nu,\Delta} = \bigcup_{i=1}^M \lambda_{i,\rho_i,\nu_i}$$

is the set of all subpaths used in the admissible optical grooming solution $\pi_{\rho,\nu,\Delta}$. The cardinality of the set is denoted $S = |\mathcal{S}_{\rho,\nu,\Delta}|$. For the sake of simplicity, we note \mathcal{S} instead of $\mathcal{S}_{\rho,\nu,\Delta}$.

$$\mathcal{A} = (a_{ij})$$

is a $\{0, 1\}^{M \times S}$ matrix; a_{ij} indicates whether SLD δ_i uses subpath \mathcal{S}_j ($a_{ij} = 1$) or not ($a_{ij} = 0$) in admissible optical grooming solution $\pi_{\rho,\nu,\Delta}$.

$$\mathcal{B} = (b_{ij})$$

is a $\{0, 1\}^{F \times S}$ arc-subpath incidence matrix; b_{ij} indicates whether arc e_i is part of subpath \mathcal{S}_j ($b_{ij} = 1$) or not ($b_{ij} = 0$).

$$\eta = \theta \cdot \tau \cdot \mathcal{A} \cdot \mathcal{B}^T = (\eta_{ij})$$

is a $\mathbb{N}^{M \times F}$ matrix; η_{ij} indicates the number of time-overlapping lightpaths on link e_j between SLD δ_i and SLDs $\delta_{i+1}, \delta_{i+2}, \dots, \delta_M$.

$$\eta^* = (\eta_j^*)_{1 \leq j \leq F}, \eta_j^* = \left\lceil \frac{1}{B} \max_{1 \leq i \leq M} \eta_{ij} \right\rceil$$

is an F -dimensional vector; η_j^* indicates the number of wavebands required on arc e_j .

$$\mathcal{IN} = \mathcal{I} \cdot \eta^*$$

is an N -dimensional vector; \mathcal{IN}_i indicates the number of input ports required in the BXC of node v_i to implement admissible optical grooming solution $\pi_{\rho, \nu, \Delta}$.

$$\mathcal{OUT} = \mathcal{O} \cdot \eta^*$$

is an N -dimensional vector; \mathcal{OUT}_i indicates the number of output ports required in the BXC of node v_i to implement admissible optical grooming solution $\pi_{\rho, \nu, \Delta}$.

$$\mathcal{G} = (g_{ij})$$

is a $\{0, 1\}^{N \times S}$ matrix; g_{ij} indicates whether vertex v_i is the source of subpath \mathcal{S}_j ($g_{ij} = 1$) or not ($g_{ij} = 0$).

$$\mathcal{H} = (h_{ij})$$

is a $\{0, 1\}^{N \times S}$ matrix; h_{ij} indicates whether vertex v_i is the termination of subpath \mathcal{S}_j ($h_{ij} = 1$) or not ($h_{ij} = 0$).

$$\mathcal{J} = \theta \cdot \tau \cdot \mathcal{A} \cdot \mathcal{G}^T = (\mathcal{J}_{ij})_{1 \leq i \leq M, 1 \leq j \leq N},$$

$$\mathcal{J}_j^* = \left\lceil \frac{1}{B} \max_{1 \leq i \leq M} \mathcal{J}_{ij} \right\rceil$$

\mathcal{J}^* is an N -dimensional vector; \mathcal{J}_j^* is the number of waveband connections added at the BXC of node v_j .

$$\mathcal{K} = \theta \cdot \tau \cdot \mathcal{A} \cdot \mathcal{H}^T = (\mathcal{K}_{ij})_{1 \leq i \leq M, 1 \leq j \leq N},$$

$$\mathcal{K}_j^* = \left\lceil \frac{1}{B} \max_{1 \leq i \leq M} \mathcal{K}_{ij} \right\rceil$$

\mathcal{K}^* is an N -dimensional vector; \mathcal{K}_j^* is the number of waveband connections dropped at the BXC of node v_j .

$$\mathcal{Q} \in \mathbb{R}_+^N,$$

$$\mathcal{Q}_i = \psi_b(\max(\mathcal{IN}_i, \mathcal{OUT}_i) + \max(\mathcal{J}_i^*, \mathcal{K}_i^*))$$

is an N -dimensional vector; \mathcal{Q}_i indicates the cost of the BXC required at node v_i to implement admissible optical grooming solution $\pi_{\rho, \nu, \Delta}$.

$$\mathcal{R} \in \mathbb{R}_+^N,$$

$$\mathcal{R}_i = \psi_w(\max(\mathcal{D}_i^*, \mathcal{E}_i^*) + B \max(\mathcal{J}_i^*, \mathcal{K}_i^*))$$

is an N -dimensional vector; \mathcal{R}_i indicates the cost of the WXC required at node v_i to implement admissible optical grooming solution $\pi_{\rho, \nu, \Delta}$. Note that the number of input/output ports, $B \max(\mathcal{J}_i^*, \mathcal{K}_i^*)$, is a multiple of the waveband size B . The parameters defining the function ψ_w may be different from those defining the function ψ_b .

The cost function $C : \Pi_{\Delta} \rightarrow \mathbb{R}_+$ is defined as:

$$C(\pi_{\rho, \nu, \Delta}) = \sum_{i=1}^N (Q_i + R_i). \tag{16.17}$$

16.5 Illustrative Examples

16.5.1 Electrical Grooming

In this subsection, we present a small example in order to highlight the reduction of network cost provided by the electrical grooming functionality. We considered the three 3 SEDs of Table 16.1. Their time diagram is shown in Fig. 16.8. To emphasize the benefit of grooming we use a fixed routing solution in which each SED is routed along its shortest path. Emitting e_1 electrical ports and receiving r_1 electrical ports are only needed to add and drop the traffic requests. Consequently, the number of such ports is independent of the routing and grooming solution. When the network does not have grooming capabilities (Fig. 16.9(a)), the SEDs δ_1 and δ_2 are routed over distinct lightpaths. Thus, we need two WDM optical channels on each of the 3–4 and the 4–7 fiber links. As a result, this routing and grooming solution requires a total of 9 WDM optical channels. Let us recall that the number of required o_1 optical ports is twice the number of WDM optical channels used. On the other hand, when the network has grooming capabilities (Fig. 16.9(b)), the SEDs δ_1 and δ_2 can be groomed together since both requests are active from 11:00 to 13:00 and they share a common path between nodes 3 and 7. Therefore, a grooming lightpath is created between these nodes using only one WDM optical channel on each of the 3–4 and the 4–7 fiber links. This grooming

Table 16.1. Characteristics of 3 SEDs

δ_i	s_i	d_i	α_i	γ_i	n_i
δ_1	2	8	8:00	14:40	$0.25 \times C_{\omega}$
δ_2	3	7	11:00	13:00	$0.25 \times C_{\omega}$
δ_3	2	6	17:00	19:30	$0.25 \times C_{\omega}$

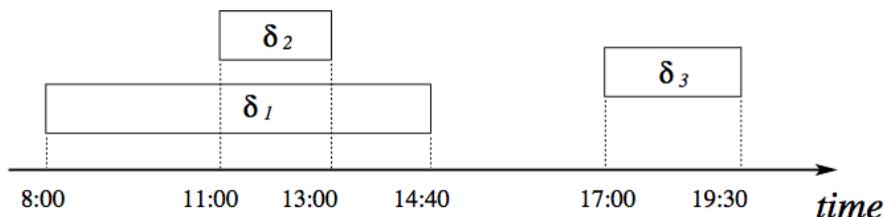


Fig. 16.8. Associated time diagram of the 3 SEDs

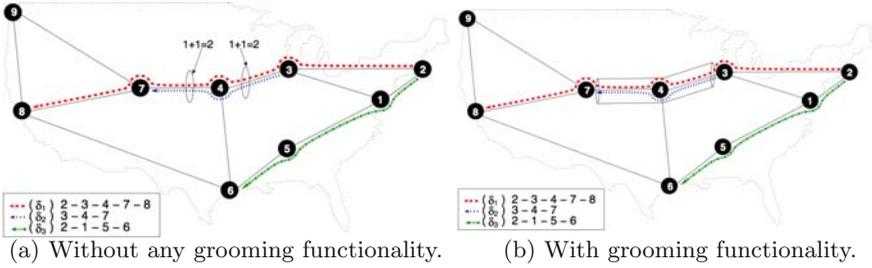


Fig. 16.9. Provisioning in a network of the 3 SEDs of Table 16.1

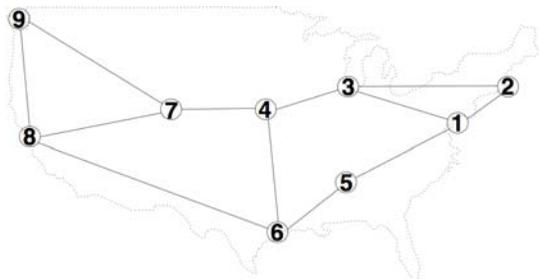
lightpath is used to carry simultaneously the two SEDs δ_1 and δ_2 between nodes 3 and 7. In our example, the SED δ_1 still needs to be routed using two additional lightpaths; the first one between nodes 2 and 3, and the second one between nodes 7 and 8. As a result, only 7 WDM optical channels are needed for this new routing and grooming solution. However, it should be noted that, in order to aggregate SEDs δ_1 and δ_2 at node 3, one additional emitting o_3 optical port and one additional receiving r_3 electrical port are needed on this node. Similarly, to de-aggregate these two SEDs at node 7, one additional emitting e_3 electrical port and one additional receiving o_3 optical port are needed. Depending on the relative costs of the different o_1 , o_3 , e_3 , and r_3 ports, the new routing and grooming solution can be more or less economical than the first solution.

16.5.2 Optical Grooming

In this subsection, we present an example to illustrate how the mathematical model of Section 16.4.2 is used to represent an instance of the optical grooming problem and to quantify the cost of a solution to this problem.

We study the problem instance of Fig. 16.10 which defines a set Δ of $M = 2$ SLDs and a graph G representing a physical network. Additionally, we choose the paths and layouts defined in Table 16.2 ($K = 2$ and $L = 3$) and define a band size of $B = 4$. For the sake of simplicity, we choose functions $\psi_w(x) = x$ and $\psi_b(x) = x$ to represent the cost of switches, i.e., we choose $a = 0$, $b = 1$ and $c = 1$ in Equation (16.16). Figure 16.11 illustrates the admissible optical grooming solution $\pi_{\rho,\nu,\Delta}$ with vectors $\rho = (1, 1)$ and $\nu = (3, 3)$ for the instance of Fig. 16.10. Path $P_{1,1}$ and layout $\lambda_{1,1,3}$ are selected for SLD δ_1 and path $P_{2,1}$ and layout $\lambda_{2,1,3}$ are selected for SLD δ_2 . The figure also shows the set of Routed Scheduled Band Groups (RSBGs) defined by the solution.

The following matrices are used to compute the cost of the chosen admissible optical grooming solution:



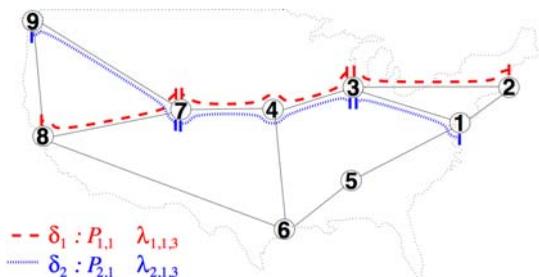
SLD	s	d	n	α	β
δ_1	2	8	8	08:00	14:00
δ_2	1	9	10	11:00	13:00

(a) Graph G representing a physical network. (b) Set Δ of $M = 2$ SLDs.

Fig. 16.10. An instance (G, Δ) of the optical grooming problem

Table 16.2. Chosen paths and layouts for the instance of Fig. 16.10

SLD	Path ($P_{i,k}$)	Layout ($\lambda_{i,k,j}$)
δ_1	$P_{1,1} = (2, 3, 4, 7, 8)$	$\lambda_{1,1,1} = \{(2, 3, 4, 7, 8)\}$
		$\lambda_{1,1,2} = \{(2, 3, 4) (4, 7, 8)\}$
		$\lambda_{1,1,3} = \{(2, 3) (3, 4, 7) (7, 8)\}$
	$P_{1,2} = (2, 1, 5, 6, 8)$	$\lambda_{1,2,1} = \{(2, 1, 5, 6, 8)\}$
		$\lambda_{1,2,2} = \{(2, 1, 5) (5, 6, 8)\}$
		$\lambda_{1,2,3} = \{(2, 1) (1, 5, 6) (6, 8)\}$
δ_2	$P_{2,1} = (1, 3, 4, 7, 9)$	$\lambda_{2,1,1} = \{(1, 3, 4, 7, 9)\}$
		$\lambda_{2,1,2} = \{(1, 3, 4) (4, 7, 9)\}$
		$\lambda_{2,1,3} = \{(1, 3) (3, 4, 7) (7, 9)\}$
	$P_{2,2} = (1, 5, 6, 8, 9)$	$\lambda_{2,2,1} = \{(1, 5, 6, 8, 9)\}$
		$\lambda_{2,2,2} = \{(1, 5, 6) (6, 8, 9)\}$
		$\lambda_{2,2,3} = \{(1, 5) (5, 6, 8) (8, 9)\}$



Subpath (λ^n)	$\Delta' \subseteq \Delta$
(1, 3)	$\{\delta_2\}$
(2, 3)	$\{\delta_1\}$
(3, 4, 7)	$\{\delta_1, \delta_2\}$
(7, 8)	$\{\delta_1\}$
(7, 9)	$\{\delta_2\}$

-- $\delta_1 : P_{1,1} \lambda_{1,1,3}$
 --- $\delta_2 : P_{2,1} \lambda_{2,1,3}$

(a) Paths and layouts.

(b) Routed Scheduled Band Groups (RSBGs).

Fig. 16.11. An admissible optical grooming solution $\pi_{\rho,\nu,\Delta}$ to the instance of Fig. 16.10

$$\theta = \begin{matrix} & \delta_1 & \delta_2 \\ \delta_1 & \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \end{matrix} \quad \gamma = \begin{matrix} & \delta_1 & \delta_2 \\ \delta_1 & \begin{pmatrix} 8 & 0 \\ 0 & 10 \end{pmatrix} \end{matrix}$$

$$\mathcal{S}_{\rho, \nu, \Delta} = \{(1, 3), (2, 3), (3, 4, 7), (7, 8), (7, 9)\}$$

$$\mathcal{A} = \begin{matrix} & s_1 & s_2 & s_3 & s_4 & s_5 \\ \delta_1 & \begin{pmatrix} 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \end{pmatrix} \end{matrix} \quad \mathcal{B} = \begin{matrix} & s_1 & s_2 & s_3 & s_4 & s_5 \\ e_{13} & \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix}$$

$$\eta = \theta \cdot \gamma \cdot \mathcal{A} \cdot \mathcal{B}^T = \begin{pmatrix} 10 & 8 & 18 & 18 & 8 & 10 \\ 10 & 0 & 10 & 10 & 0 & 10 \end{pmatrix} \quad \eta^* = \begin{pmatrix} 3 \\ 2 \\ 5 \\ 5 \\ 2 \\ 3 \end{pmatrix}$$

$$\mathcal{I} = \begin{matrix} & e_{13} & e_{23} & e_{34} & e_{47} & e_{78} & e_{79} \\ v_1 & \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix} \quad \mathcal{O} = \begin{matrix} & e_{13} & e_{23} & e_{34} & e_{47} & e_{78} & e_{79} \\ \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \end{matrix}$$

$$\mathcal{IN} = \mathcal{I} \cdot \eta^* = (0, 0, 5, 5, 0, 0, 5, 2, 3)$$

$$\mathcal{OUT} = \mathcal{O} \cdot \eta^* = (3, 2, 5, 5, 0, 0, 5, 0, 0)$$

$$\mathcal{G} = \begin{matrix} & s_1 & s_2 & s_3 & s_4 & s_5 \\ v_1 & \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \end{matrix} \quad \mathcal{H} = \begin{matrix} & s_1 & s_2 & s_3 & s_4 & s_5 \\ v_1 & \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix}$$

$$\begin{aligned}
 \mathcal{T} &= \begin{matrix} & v_1 & v_2 & v_3 & \dots & v_9 \\ \delta_1 & \left(\begin{array}{cccccc} 0 & 1 & 0 & \dots & 0 \\ 1 & 0 & 0 & \dots & 0 \end{array} \right) & \mathcal{U} &= \begin{matrix} & v_1 & \dots & v_7 & v_8 & v_9 \\ \delta_1 & \left(\begin{array}{cccccc} 0 & \dots & 0 & 1 & 0 \\ 0 & \dots & 0 & 0 & 1 \end{array} \right) \end{matrix} \\
 \mathcal{D} &= \theta \cdot \gamma \cdot \mathcal{T} = \begin{pmatrix} 10 & 8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 10 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} & \mathcal{D}^* &= (10 \ 8 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0) \\
 \mathcal{E} &= \theta \cdot \gamma \cdot \mathcal{U} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 8 & 10 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 10 \end{pmatrix} & \mathcal{E}^* &= (0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 8 \ 10) \\
 \mathcal{J} &= \theta \cdot \gamma \cdot \mathcal{A} \cdot \mathcal{G}^T = \begin{pmatrix} 10 & 8 & 18 & 0 & 0 & 0 & 18 & 0 & 0 \\ 10 & 0 & 10 & 0 & 0 & 0 & 10 & 0 & 0 \end{pmatrix} & \mathcal{J}^* &= (3 \ 2 \ 5 \ 0 \ 0 \ 0 \ 5 \ 0 \ 0) \\
 \mathcal{K} &= \theta \cdot \gamma \cdot \mathcal{A} \cdot \mathcal{H}^T = \begin{pmatrix} 0 & 0 & 18 & 0 & 0 & 0 & 18 & 8 & 10 \\ 0 & 0 & 10 & 0 & 0 & 0 & 10 & 0 & 10 \end{pmatrix} & \mathcal{K}^* &= (0 \ 0 \ 5 \ 0 \ 0 \ 0 \ 5 \ 2 \ 3) \\
 \mathcal{Q} &= (6 \ 4 \ 10 \ 5 \ 0 \ 0 \ 10 \ 4 \ 6) & \mathcal{R} &= (22 \ 16 \ 20 \ 0 \ 0 \ 0 \ 20 \ 16 \ 22)
 \end{aligned}$$

The cost of the admissible optical grooming solution $\pi_{\rho,\nu,\Delta}$ according to (16.17) is:

$$C(\pi_{\rho,\nu,\Delta}) = 28 + 20 + 30 + 5 + 0 + 0 + 30 + 20 + 28 = 161. \tag{16.18}$$

16.6 Conclusion

This chapter presented mathematical formulations to quantify the cost of provisioning scheduled connections in multi-layer optical networks. The specific cases of electrical grooming in EXC/WXC networks and optical grooming in WXC/BXC networks were considered. The formulations may serve as a basis to assess the costs and benefits of introducing an additional network layer in an existing network, or to determine the impact on the network-wide cost of a change in the technologies used to implement a cross-connect. The use of scheduled connections make it possible to take into account dynamic changes on the traffic load in the cost/benefit analysis of multi-layer networks.

The formulations could be extended to compare the impact on network cost of alternative resiliency approaches such as link or path protection, dedicated or shared protection, etc. Another extension could be the consideration of the uncertainty about the tear-down date of SEDs or SLDs, since connections terminating later than their expected tear-down date might be seizing network resources allocated to demands that were originally time-disjoint with respect to the seizing connection. It is reasonable to expect that, the longer a connection lasts in the network, the higher the uncertainty about its actual tear-down date.

References

1. J. Kuri, N. Puech, M. Gagnaire, E. Dotaro and R. Douville, "Routing and Wavelength Assignment of Scheduled Lightpath Demands," *IEEE Journal on Selected Areas in Communications*, Vol. 21, N. 8, pp. 1231–1240, October 2003.
2. J. Kuri, "Optimization Problems in WDM Optical Transport Networks with Scheduled Lightpath Demands," *Ph.D. Thesis*, Ecole Nationale Supérieure des Télécommunications. Paris, France. September 2003.
3. N. Skorin-Kapov, "Heuristic algorithms for the routing and wavelength assignment of scheduled lightpath demands in optical networks," *IEEE Journal on Selected Areas in Communications*, Vol. 24, NO. 8, August 2006.
4. H. G. Ahn, T.-J. Lee, M. Y. Chung and H. Choo, "RWA on Scheduled Lightpath Demands in WDM Optical Transport Networks with Time Disjoint Paths," *Proceedings of the International Conference on Information Networking (ICOIN 2005)*, Jeju, Korea, February 2005.
5. C. V. Saradhi and M. Gurusami, "Graph Theoretic Approaches for Routing and Wavelength Assignment of Scheduled Lightpath Demands in WDM Optical Networks," *Proceedings of the 2nd International Conference on Broadband Networks (Broadnets 2005)*, Boston, MA, October 2005.
6. J. Kuri, N. Puech, M. Gagnaire and E. Dotaro, "Routing and Wavelength Assignment of Scheduled Lightpath Demands in a WDM Optical Transport Network," *Proceedings of the First International Conference in Optical Communications and Networks (ICOON 2002)*, Singapur, November 2002.
7. T. Li and B. Wang, "On optimal survivability design in WDM optical networks under a scheduled traffic model," *Proceedings of the 5th International Workshop on Design of Reliable Communication Networks (DRCN 2005)*, Ischia, Italy, October 2005.
8. T. Li and B. Wang, "On Optimal Survivability Design under a Scheduled Traffic Model in Wavelength-Routed Optical Mesh Networks," *Proceedings of the 4th Annual Communication Networks and Services research Conference (CNSR 2006)*, Moncton, Canada, May 2006.
9. C. V. Saradhi, L. K. Wei and M. Gurusami, "Provisioning Fault-Tolerant Scheduled Lightpath Demands in WDM Mesh Networks," *Proceedings of the First International Conference on Broadband Networks (BROADNETS 2004)*, San Jose, CA, October 2004.
10. E. A. Doumith and M. Gagnaire, "Traffic Routing in Multi-Layer Optical Network Considering Rerouting and Grooming Strategies," *Proceedings of the IEEE Global Communications Conference (Globecom)*, San Francisco, CA, December 2006.
11. E. A. Doumith, M. Gagnaire, O. Audouin, and R. Douville, "Network Nodes Dimensioning Assuming Electrical Traffic Grooming in an Hybrid OXC/EXC WDM Network," *Proceedings of the 2nd IEEE/Create-Net International Conference on Broadband Networks (BroadNets'05)*, Boston, MA, 2005.
12. E. A. Doumith and M. Gagnaire, "Traffic Grooming in Multi-Layer WDM Networks: Meta-heuristics versus Sequential Algorithms," *Proceedings of the IEEE Eunice Open European Summer School*, Stuttgart, Germany, September 2006.

13. E. A. Doumith, M. Gagnaire, O. Audouin and R. Douville, "From Network Planning to Traffic Engineering for Optical VPN and Multi-Granular Random Demands," *Proceedings of the 2nd IEEE/Create-Net International Conference on Broadband Networks (BroadNets'05)*, Phoenix, AZ, April 2006.
14. D. Eppstein, "Finding the k Shortest Paths," *IEEE Symposium on Foundations of Computer Science*, pp. 154–165, 1994.

All-Optical Traffic Grooming

Volkan Kaman, Henrik N. Poulsen, Roger J. Helkey, John E. Bowers

17.1 Introduction

Next-generation optical communication networks are expected to exceed aggregate capacities of hundreds of terabits per second to support the recent growth in commercial broadband access fueled by applications such as Video-on-Demand (VoD) and Voice-over-IP (VoIP), and to meet the high-performance demands of grid-computing networks for research consortiums and military applications [15, 26, 31, 32]. The continuous increase in need for higher bandwidth has also necessitated in parallel a continuous reduction in cost per bit, which has paved the way for the all-optical revolution. As higher bandwidth with smaller footprint and lower power consumption scales better with lower cost optics, the past 30 years have seen electronics get gradually pushed out of the core communication network to the lower bandwidth edge user interfaces. The advantage of transitioning to all-optical communication is most evident first with the introduction of the erbium-doped fiber amplifier (EDFA) in the early 1990s. The EDFA not only increased the all-optical reach of bandwidth independent signals by eliminating costly optical-electrical-optical (OEO) repeaters, but it also revolutionized point-to-point communication system capacity by allowing widespread implementation of dense wavelength division multiplexing (DWDM). The second confirmation of the benefit of all-optical communication is the introduction of optical add-drop multiplexers (OADM) at metro and backbone network nodes in the early 2000s. Similarly to the EDFA, the OADM allowed the removal of costly in-line OEO transponders by providing all-optical bypass of express traffic. Furthermore, in realizing the concept of the all-optical core ring network, the OADM provides significant optics savings by driving the grooming electronics switch to the edge of the network for sub-wavelength processing of the local add-drop traffic.

While the EDFA and OADM have allowed for the realization of today's all-optical point-to-point systems and core ring networks, the current shift to higher data rates of 40 Gb/s SONET, 100 Gb Ethernet and beyond opens up the possibility for further cost-saving all-optical innovations in high-capacity

edge and core networks. This chapter reviews the next-generation all-optical traffic grooming technologies for both the core and edge networks. Section 17.2 reviews the state-of-the-art smart optical switching technologies for all-optical wavelength grooming in the core network as these networks evolve to interconnected rings and mesh architectures. Section 17.3 reviews all-optical circuit grooming, or optical time division multiplexing (OTDM), for aggregating low-speed sub-wavelength circuits into wavelengths between the core and edge networks. Finally, Section 17.4 reviews all-optical edge grooming techniques such as optical packet switching (OPS) and optical burst switching (OBS) for reducing costly electronics in edge grooming switches and IP routers, and Section 17.5 concludes the chapter.

17.2 All-Optical Wavelength Grooming

The aggregate bandwidth capacity increase in next-generation metro and long haul core optical networks requires a considerable increase in DWDM fiber link capacity, which is achieved by increasing both the number of wavelengths and the data rate per wavelength using spectrally efficient modulation formats as recently demonstrated with a 25.6 Tb/s transmission experiment [6]. These high-capacity DWDM links are interconnected into dynamically reconfigurable all-optical mesh networks with efficient wavelength provisioning and optimized use of network elements for both capital and operational cost savings. This agile photonic network is enabled by the next-generation of remotely reconfigurable OADM (ROADM) with express traffic switching and local add-drop wavelength grooming capabilities. The ROADM should support multiple degree nodes, which is defined as the number of DWDM fiber network ports terminating at the node. It should also allow each add-drop port to have fully flexible access to the network to ensure efficient network utility of each access transponder. The latter is of high significance in a multi-degree ROADM design as transponders and regenerators dominate the overall core network cost and should be minimized to the lowest possible required by the network [5, 31]. The key enablers for all-optical wavelength grooming of core network traffic are the dramatic savings in (i) cost per express wavelength switching in the optical domain by removing in-line OEO transponders and electronic switches, and (ii) cost per access wavelength switching by removing redundant edge OEO transponders through flexible optical add-drop.

Traditional degree-two nodes in ring networks have relied upon static OADMs that can transit and add-drop only fixed wavelengths without reconfigurability. These OADMs are either based on a three-port WDM filter approach (Fig. 17.1a) or a pair of concatenated optical demultiplexers (DMUX) and multiplexers (MUX) for splitting and recombining all DWDM wavelengths (Fig. 17.1b). Next-generation degree-two ROADMs are based on wavelength blockers (WB) using liquid crystals (LC), or one-dimensional micro-electro-mechanical-system (1-D MEMS) switches, or integrated planar

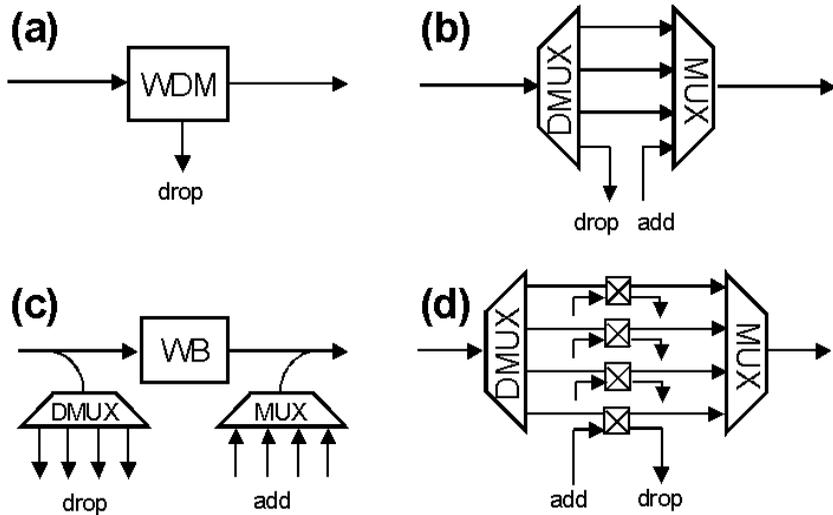


Fig. 17.1. Schematics of degree-two static OADM's based on (a) a WDM filter, (b) concatenated DMUX and MUX; and ROADM's based on (c) WB's, and (d) concatenated DMUX and MUX with 2×2 switches for add-drop access

lightwave circuits (PLC), where all techniques process and power balance express traffic at a per wavelength basis [18, 23]. Add-drop access for both 1-D MEMS and LC-based ROADMs is achieved by tapping the aggregate DWDM signal before and after the WB and separating the channels using fixed wavelength D/MUXs (Fig. 17.1c). The PLC-based ROADM has the advantage of utilizing the already integrated express D/MUXs for fixed wavelength add-drop access by further integrating per channel 2×2 switches onto the PLC device (Fig. 17.1d). While degree-two ROADMs provide reconfigurable all-optical wavelength grooming, fully flexible add-drop access requires the use of additional tunable filters and optical switches, the cost of which can be prohibitive for high add-drop ratios.

As next-generation networks evolve into interconnected rings and mesh architectures, multi-degree ROADM nodes with flexible add-drop capability gain high significance. The add-drop flexibility criteria that eliminate additionally redundant and expensive electronics with cheap optics are summarized as follows [10]:

1. *directionless access* with any add-drop port to any DWDM network port connectivity,
2. *colorless access* with any wavelength-transparent add-drop port to any DWDM network port connectivity, and
3. *contentionless access* between all same-wavelength add-drop ports.

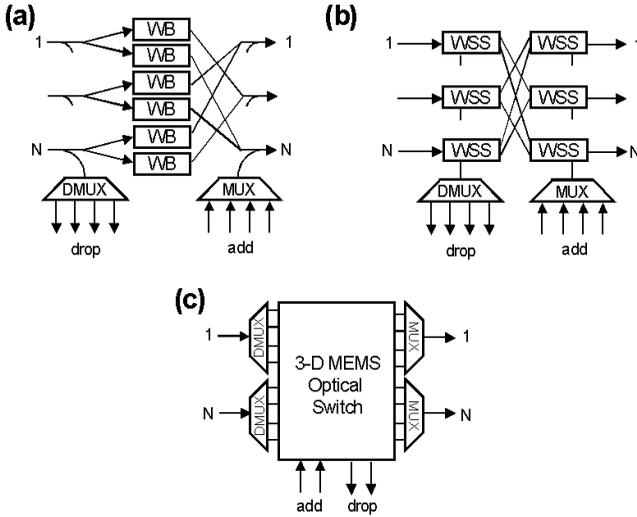


Fig. 17.2. Schematics of multi-degree ROADM architectures (a) broadcast-and-select using WBs, (b) optical cross-connect using WSSs, (c) WSXC using 3-D MEMS optical switch

Multi-degree ROADMs can be realized using three distinct technologies as summarized in Fig. 17.2: (a) a broadcast-and-select architecture using WBs, (b) an optical cross-connect (or broadcast-and-select) architecture using wavelength selective switches (WSS), and (c) a wavelength-selective cross-connect (WSXC) architecture using a large-scale 3-D MEMS optical switch.

The broadcast-and-select architecture shown in Fig. 17.2a is based on the same 1-D MEMS or LC WB technologies as discussed for degree-two ROADMs [30]. While passive optical splitters are used to broadcast and combine each incoming and outgoing DWDM network port, the WBs select the desired wavelengths to be transmitted. An inherent disadvantage of this architecture is that as the degree of the node N increases, the required number of WBs increases quadratically to $N(N - 1)$ while the ROADM express through loss increases significantly as well. This multi-degree ROADM also physically separates the express wavelength switching from the add-drop switching by tapping the DWDM signals before and after the core express switch. While the degree based modularity of this ROADM is attractive from a network restoration perspective, this also severely complicates the ability to provide flexible add-drop access with the aforementioned flexibility criteria.

A more attractive alternative to WB's is to use 1-D MEMS or LC based $1 : K$ WSS devices (typically with $K < 10$), which distribute an incoming DWDM signal into any of the desired K outputs [2, 20]. A multi-degree ROADM node can be achieved by cascading two WSS devices per network port and inter-connecting $N - 1$ WSS ports from one network port to the others

as shown in Fig. 17.2b [1]. While any of the remaining $K - N + 1$ WSS ports can be used for colorless local add-drop access of a single transponder, the limited number of available WSS ports requires optical aggregation (using D/MUX, splitters, and EDFAs) of the transponders before accessing the WSS. Furthermore, directionless and wavelength contentionless add-drop access requires further optical processing at the edge, which becomes undesirable for high-degree nodes with high add-drop ratios [13]. While the preceding described the optical cross-connect architecture using $2N$ WSS devices with a fixed through loss, typical ROADM implementations use the broadcast-and-select architecture by cascading an input splitter with a single WSS device per network port [1]. This architecture has the advantage that the number of required WSS devices per node is reduced to N ; however, the challenges with flexible add-drop access remain with an additional express insertion loss drawback.

A multi-degree ROADM can also be implemented with a WSXC architecture using a large-scale 3-D MEMS optical switch combined with a set of D/MUXs (Fig. 17.2c) [14]. While non-blocking 3-D MEMS switches with low optical insertion losses are commercially available with 320 ports, these matrix switches have also been shown to scale beyond 1,000 ports [9, 17]. Wavelength switching between DWDM network ports is achieved by directing each wavelength separately to a MEMS port through the use of modular or integrated D/MUX at the DWDM network ports. Furthermore, fully agile add-drop access is realized due to the colorless and non-blocking nature of each MEMS port. A transponder or regenerator can connect to the entire core network at any wavelength using a single MEMS port. Since the add-drop ports use the already existing express filtering devices, no additional elements are required for the add-drops, reducing cost compared to the preceding WSS architectures.

The three multi-degree ROADM technologies are compared in Fig. 17.3 in terms of WSS functionality and evolution. The first generation WSS is the WB with a single input and output. The limited number of ports on this 1:1 WSS not only results in a quadratic need of these devices in implementing a multi-degree ROADM, but also the absence of local access ports requires additional components with limited add-drop flexibility. The next-generation device is the 1: K WSS, which has a single input port and several output ports. The number of devices required to implement a multi-degree ROADM scales linearly with the number of network ports. While a limited number of the K ports can be used for colorless local access, full add-drop flexibility and high add-drop ratios requires additional components and complicated access architectures. Finally, the integrated WSXC realizes the multi-degree ROADM through a $N:N$ WSS device with additional ports available for fully flexible add and drop access at wavelength granularity with simple operational management. While the latter approach provides the most flexible multi-degree ROADM due to the non-blocking switching nature of each wavelength port, the former two architectures provide a modular growth possibility with less add-drop flexibility. While modular growth of 3-D MEMS switches has been proposed, it can be concluded that network port modularity and

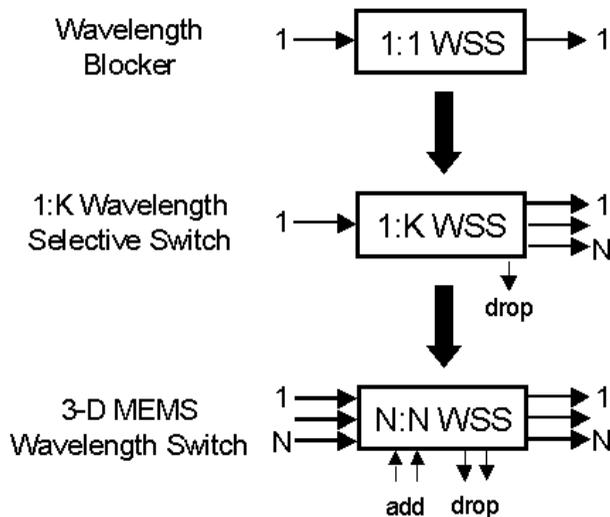


Fig. 17.3. Comparison of multi-degree ROADMs technologies in terms of WSS functionality and evolution

local add-drop flexibility of a multi-degree ROADMs are competing design criteria [13].

17.3 All-Optical Circuit Grooming

While DWDM is widely utilized to upgrade fiber bandwidth capacity, the cost, provisioning, and management of a huge number of wavelength channels can be highly problematic when considering dynamic mesh networks as opposed to point-to-point systems. Increasing the bit rate of each wavelength channel using traditional serial TDM circuit grooming while keeping the number of DWDM wavelengths to a cost effective and reasonably manageable number is therefore required to satisfy the bandwidth requirements of next-generation high-capacity networks. While TDM based on electronic grooming is currently being deployed at 40 Gb/s in commercial networks and demonstrated at 100 Gb/s in lab trials [19, 29], the speed and cost limitations of electronics will most likely require all-optical circuit grooming or OTDM techniques for achieving higher data rates of 160 Gb/s and beyond per wavelength [33]. As sub-wavelength multiplexing and demultiplexing are performed in the optical domain, OTDM systems have the advantage of needing opto-electronic component bandwidths only at the lower speed tributary rate. For example, a 160 Gb/s data rate can be achieved with commercially available low-cost 10 GHz devices [24]. Optical grooming or aggregation capability between the edge and core networks also allows for efficient utilization of resources for dynamic

sub-wavelength traffic. Furthermore, as lower speed circuits within a given wavelength may have different destinations, the ability to drop the desired circuits and insert new circuits in the available slots, or add-drop multiplexing, becomes a desirable functionality [27]. While traditional TDM requires electronic processing of all sub-wavelength traffic within a given wavelength, optical add-drop multiplexing using OTDM techniques can provide complete optical transparency to express sub-wavelength traffic from source to destination.

The basic principle of OTDM is illustrated in a point-to-point system in Fig. 17.4a. A typical OTDM transmitter consists of an optical pulse source that emits short pulses at the desired wavelength with a repetition rate of T , which is determined by the bit rate B at which electronics can follow. A passive $1:M$ splitter broadcasts these pulses into separate paths, where each pulse stream is independently data modulated and time delayed. All M pulse streams are then bit interleaved into a single serial data stream with an aggregate data rate of $B \times M$. The number of serial channels that can be packed into a single wavelength depends on the width of the optical pulses. On the OTDM receiver side, a $1:M$ splitter broadcasts the aggregate data stream to each of the M tributary detectors. A phase-locked optical gating device (DMUX) is utilized prior to the detector for extracting the desired sub-wavelength channel. The concept of add-drop multiplexing using the aforementioned OTDM techniques is shown in Fig. 17.4b. An optical gate extracts the desired channel to be dropped at the node while the rest of the channels are transmitted transparently. A new channel is then inserted in the available time slot.

While various OTDM technologies have been demonstrated in lab trials based on all-optical fiber techniques for ultra-high-speed single wavelength data rates [25, 33], the high number of devices required in OTDM sub-systems

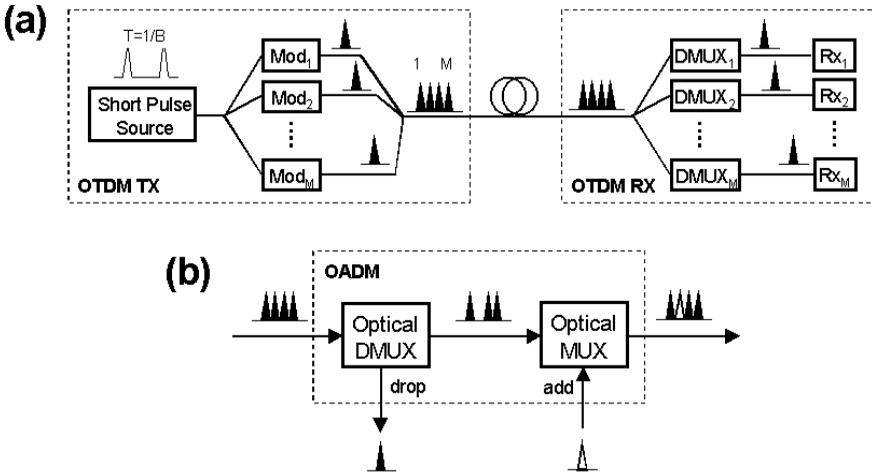


Fig. 17.4. (a) Basic schematic of an OTDM transmitter and receiver in a point-to-point system, (b) concept of OTDM optical add-drop multiplexing

makes photonic integration essential for commercial viability. Compact optical short pulse sources using a tandem of integrated electro-absorption modulators [8] and further integration with semiconductor optical amplifiers (SOA) for compensating losses have been demonstrated [22]. Further chip level integration of splitters, arrays of modulators and amplifiers on PLC platforms also show promise for commercial viability [16]. An OTDM demultiplexing receiver with an integrated electro-absorption modulator for optical gating followed by a detector has been demonstrated at 40 Gb/s [7]. Further integration of optical amplification has also been achieved for better receiver sensitivity performance [21]. Monolithically integrated SOA devices have been used for add-drop applications [28] while a single electro-absorption modulator has also been shown to simultaneously demultiplex and detect a single channel while transparently transmitting the express channels [12]. Using this capability, an alternative OTDM receiver architecture that imitates high-speed electrical receiver sub-systems with the benefit of operating at the tributary base rate has also been proposed [11]. The compact OTDM receiver eliminates the $1:M$ splitter and requires a single input fiber into a series of integrated and cascaded modulators to achieve high-speed serial-optical to direct tributary-speed parallel-electrical conversion. Apart from photonic integration, there are several other challenges that need to be considered for OTDM to have commercial penetration. The generation, multiplexing, and gating of such short pulses imposes very strict and challenging polarization, crosstalk, extinction ratio and timing with active signal processing requirements. Finally, further challenges of optical transmission of ultra-high data rates in dynamic networks will require thoughtful consideration.

17.4 All-Optical Edge Grooming

Present day networks are based on optical circuit switching with dedicated traffic paths carrying traffic from source to destination. As network bandwidth demands increase, DWDM systems with higher count channels and faster data rates are utilized to provide the required bandwidth capacity. While DWDM is a very effective means of using the bandwidth of the installed fiber base, such switching paradigms exhibit poor performance when sub-wavelength traffic is considered. As described in the preceding section, all-optical circuit grooming using OTDM techniques can help in aggregating several optical channels onto one wavelength. However, the limited granularity and static nature of OTDM switching does not render this as an efficient solution in edge networks. Therefore, traffic grooming at the edge of a network, which presently makes use of the fine granularity electronics can offer, is becoming increasingly important to improve the usage of the available optical resources and to increase the overall network utilization. In SONET and SDH networks, this is mainly accomplished by statistical multiplexing, where data from a large number of logical channels are carried on a single wavelength in the

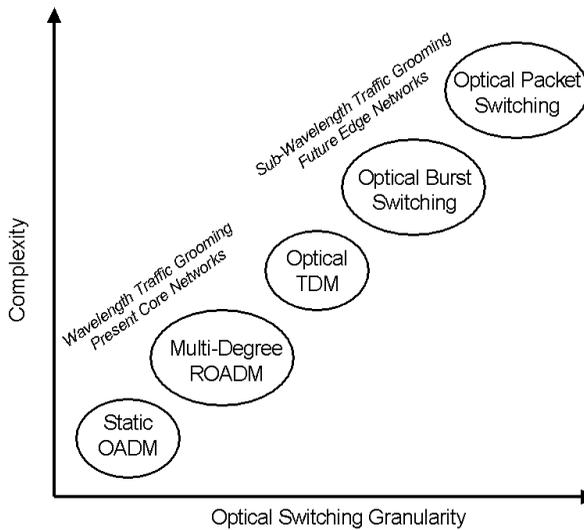


Fig. 17.5. Complexity of the various all-optical traffic grooming technologies

physical medium. The available bandwidth is dynamically allocated only towards active channels, which enables more devices to be connected than other multiplexing techniques. However, as services such as VoD and VoIP are being introduced, the requirements for low latency and high granularity are increasing. Due to the high cost and power consumption of high bit-rate electronics and with progressively more emphasis on transparency, all-optical edge grooming using OPS and OBS are becoming more and more interesting [34].

From a traffic utilization point of view, the highest gains are currently envisaged by using OPS since it can provide both the granularity and latency required for multi-media services [3, 35]. Instead of encapsulating each IP packet into larger SONET frames that are transmitted in a synchronous bit stream, with OPS each IP packet is transmitted by itself with an additional small header. Thus, IP packets arrive at the end destination asynchronously as they are not electronically processed and synchronized at each node in the network. While OPS provides highly efficient and granular switching, it also requires new technology at the edge that also permeates into the whole core network. The current lack of switches with nanosecond switching times and optical buffers are prohibitive factors for OPS networks with substantial development time required.

A hybrid between optical circuit and packet switching is OBS [4]. In contrast to OPS where packets are sent out one by one, OBS networks aggregate packets with the same end destination into larger packets, which are denoted as bursts. With significantly larger packets, the speed requirements of the switching elements in the core network are a lot less strict at the expense of decreased network utilization. The minimum length of a burst is determined

by the switching technology used and has an upper limit for the time allowed in which a burst can be assembled. As a burst cannot arbitrarily wait long enough for IP packets to arrive in assembling the burst, many bursts are likely to be timed out and transmitted without the minimum number of IP packets, in which case the network utilization is decreased. While OBS alleviates the issue of fast switching, it still shares the issue of optical buffers as in OPS.

The complexity of the various technologies discussed in this chapter for all-optical traffic grooming is summarized in Fig. 17.5 with respect to switching granularity.

17.5 Conclusion

In the quest for high-capacity and low-cost optical networks, traditional electronic traffic grooming is being replaced by more efficient all-optical technologies. Multi-degree ROADMs with wavelength grooming capabilities are already being deployed in commercial core networks while active research in OTDM, OPS, and OBS is underway for all-optical sub-wavelength grooming of edge traffic.

References

1. E.B. Basch, R. Egorov, S. Gringeri, and S. Elby. Architectural tradeoffs for reconfigurable dense wavelength-division multiplexing systems. *IEEE Journal in Selected Topics on Quantum Electronics*, 12:615–626, 2006.
2. G. Baxter, S. Frisken, D. Abakoumov, Z. Hao, I. Clarke, A. Bartos, and S. Poole. Highly programmable wavelength selective switch based on liquid crystal on silicon switching elements. In *Optical Fiber Communications Conference (OFC/NFOEC), paper OTuF2*, 2006.
3. D.J. Blumenthal, J.E. Bowers, L. Rao, H.-F. Chou, S. Rangarajan, and H.N. Poulsen. Optical signal processing for optical packet switching networks. *IEEE Communications Magazine*, 41:S23–S29, 2003.
4. Y. Chen, C. Qiao, and X. Yu. Optical burst switching: A new area in optical networking research. *IEEE Network Magazine*, 18(1):16–23, 2004.
5. O. Gerstel and H. Raza. Predeployment of resources in agile photonic networks. *IEEE Journal of Lightwave Technology*, 22:2236–2244, 2004.
6. A.H. Gnauck, G. Charlet, P. Tran, P. Winzer, C. Doerr, J. Centanni, E. Burrows, T. Kawanishi, T. Sakamoto, and K. Higuma. 25.6-Tb/s C+L-band transmission of polarization-multiplexed RZ-DQPSK signals. In *Optical Fiber Communications Conference (OFC/NFOEC), paper PDP19*, 2007.
7. V. Kaman, Y.J. Chiu, T. Liljeberg, S.Z. Zhang, and J.E. Bowers. A compact 40 Gbit/s demultiplexing receiver based on integrated tandem electroabsorption modulators. *IEEE Electronics Letters*, 36:1943–1944, 2000.
8. V. Kaman, Y.J. Chiu, T. Liljeberg, S.Z. Zhang, and J.E. Bowers. Integrated tandem traveling-wave electroabsorption modulators for > 100 Gbit/s OTDM applications. *IEEE Photonics Technology Letters*, 12:1471–1473, 2000.

9. V. Kaman, R.J. Helkey, and J.E. Bowers. Compact and scalable 3-D MEMS optical switches. *Journal of Optical Networks*, 6:19–24, 2007.
10. V. Kaman, R.J. Helkey, and J.E. Bowers. Multi-degree ROADMs with agile add-drop access. In *Photonics in Switching Conference, paper TuA2*, 2007.
11. V. Kaman, A.J. Keating, S.Z. Zhang, and J.E. Bowers. Electroabsorption modulator as a compact OTDM demultiplexing receiver. In *ECOC, paper 9.4.6*, 2000.
12. V. Kaman, A.J. Keating, S.Z. Zhang, and J.E. Bowers. Simultaneous OTDM demultiplexing and detection using an electroabsorption modulator. *IEEE Photonics Technology Letters*, 12:711–713, 2000.
13. V. Kaman, S. Yuan, O. Jerphagnon, R.J. Helkey, and J.E. Bowers. Comparison of wavelength-selective cross-connect architectures for reconfigurable all-optical networks. In *Photonics in Switching Conference, paper 14.1*, 2006.
14. V. Kaman, X. Zheng, S. Yuan, J. Klingshirn, C. Pusarla, R.J. Helkey, O. Jerphagnon, and J.E. Bowers. A 32×10 gbit/s DWDM metropolitan network demonstration using wavelength-selective photonic cross-connects and narrow-band EDFAs. *IEEE Photonics Technology Letters*, 17:1977–1979, 2005.
15. G. Karmous-Edwards. Today's optical network research infrastructures for E-Science applications. In *Optical Fiber Communications Conference (OFC/NFOEC), paper OWU3*, 2006.
16. K. Kato and Y. Tohmori. PLC hybrid integration technology and its application to photonic components. *IEEE Journal in Selected Topics on Quantum Electronics*, 6(1):4–13, 2000.
17. J. Kim, C.J. Nuzman, B. Kumar, D.F. Liewwen, J.S. Kraus, A. Weiss, C.P. Lichtenwalner, A.R. Papazian, R.E. Frahm, N.R. Basavanhally, D.A. Ramsey, V.A. Aksyuk, F. Pardo, M.E. Simon, V. Lifton, H.B. Chan, M. Haueis, A. Gasparyan, H.R. Shea, S. Arney, C.A. Bolle, P.R. Kolodner, R. Ryf, D.T. Neilson, and J. V Gates. 1100×1100 port MEMS-based optical crossconnect with 4-dB maximum loss. *IEEE Photonics Technology Letters*, 15:1537–1539, 2003.
18. J. Kondis, B.A. Scott, A. Ranalli, and R. Lindquist. Liquid crystals in bulk optics-based DWDM optical switches and spectral equalizers. In *IEEE LEOS Annual Meeting*, 1; 292–293, 2001.
19. E. Lach and K. Schuh. Recent advances in ultrahigh bit rate ETDM transmission systems. *IEEE Journal of Lightwave Technology*, 24:4455–4467, 2006.
20. D.M. Marom, D.T. Neilson, D.S. Greywall, C.-S. Pai, N.R. Basavanhally, V.A. Aksyuk, D.O. Lopez, F. Pardo, M.E. Simon, Y. Low, P. Kolodner, and C. A. Bolle. Wavelength-selective $1 \times k$ switches using free-space optics and MEMS micromirrors: Theory, design, and implementation. *IEEE Journal of Lightwave Technology*, 23:1620–1630, 2005.
21. B. Mason, J.M. Geary, J.M. Freund, A. Ougazzaden, C. Lentz, K. Glogovsky, G. Przybylek, L. Peticolas, F. Walters, L. Reynolds, J. Boardman, T. Kercher, M. Rader, D. Monroe, and L. Ketelsen. 40 Gb/s photonic integrated receiver with -17 dBm sensitivity. In *Optical Fiber Communications Conference (OFC), paper FB10*, 2002.
22. B. Mason, A. Ougazzaden, C.W. Lentz, K.G. Glogovsky, C.L. Reynolds, G.J. Przybylek, R.E. Leibenguth, T.L. Kercher, J.W. Boardman, M.T. Rader, J.M. Geary, F.S. Walters, L.J. Peticolas, J.M. Freund, S.N.G. Chu, A. Sirenko, R.J. Jurchenko, M.S. Hybertsen, L.J.P. Ketelsen, and G. Raybon. 40-Gb/s tandem electroabsorption modulator. *IEEE Photonics Technology Letters*, 14(1):27–29, 2002.

23. S. Mechels, L. Muller, D. Morley, and D. Tillett. 1D MEMS-based wavelength switching subsystem. *IEEE Communications Magazine*, 41:88–94, 2003.
24. B. Mikkelsen, G. Raybon, R.J. Essiambre, K. Dreyer, Y. Su, J.E. Johnson, G. Shtengel, A. Bond, D.G. Moodie, and A.D. Ellis. 160 Gbit/s single-channel transmission over 300 km nonzero-dispersion fiber with semiconductor based transmitter and demultiplexer. In *ECOC, paper PD 2-3*, 1999.
25. M. Nakazawa, T. Yamamoto, and K.R. Tamura. 1.28 Tbit/s - 70 km OTDM transmission using third- and fourth-order simultaneous dispersion compensation with a phase modulator. In *ECOC, paper PD 2.6*, 2000.
26. M.J. O’Mahony, C. Politi, D. Klonidis, R. Nejabati, and D. Simeonidou. Future optical networks. *IEEE Journal of Lightwave Technology*, 24:4684–4696, 2006.
27. S. Pau, J. Yu, K. Kojima, N. Chand, and V. Swaminathan. 160-Gb/s all-optical MEMS time-slot switch for OTDM and WDM applications. *IEEE Photonics Technology Letters*, 13:1460–1462, 2001.
28. H.N. Poulsen, K.S. Jepsen, A.T. Clausen, A. Buxens, K.E. Stubkjaeri, R. Hess, M. Dulk, and G. Melchior. Fast optical signal processing in high bit rate OTDM systems. In *IEEE LEOS Annual Meeting*, volume 1, pages 67–68, 1998.
29. G. Raybon, P.J. Winzer, and C.R. Doerr. 1-Tb/s (10×107 Gb/s) electronically multiplexed optical signal generation and WDM transmission. *IEEE Journal of Lightwave Technology*, 25:233–238, 2007.
30. J.-K. Rhee, I. Tomkos, and M.-J. Li. A broadcast-and-select OADM optical network with dedicated optical-channel protection. *IEEE Journal of Lightwave Technology*, 21(1):25–31, 2003.
31. A.A.M. Saleh and J. M. Simmons. Evolution toward the next-generation core optical network. *IEEE Journal of Lightwave Technology*, 24:3303–3321, 2006.
32. R.E. Wagner, J.R. Igel, R. Whitman, M.D. Vaughn, A. B. Ruffin, and S. Bickham. Fiber-based broadband-access deployment in the United States. *IEEE Journal of Lightwave Technology*, 24:4526–4540, 2006.
33. H.G. Weber, R. Ludwig, S. Ferber, C. Schmidt-Langhorst, M. Kroh, V. Marembert, C. Boerner, and C. Schubert. Ultrahigh-speed OTDM-transmission technology. *IEEE Journal of Lightwave Technology*, 24:4616–4627, 2006.
34. S.J.B. Yoo. Optical packet and burst switching technologies for the future photonic internet. *IEEE Journal of Lightwave Technology*, 24:4468–4492, 2006.
35. M. Zirngibl. IRIS: Optical switching technologies for scalable data networks. In *Optical Fiber Communications Conference (OFC/NFOEC), paper OWP1*, 2006.

Index

- admissible solution, 268, 270, 271, 273, 274, 276
- aggregated demand, 257
- all-optical circuit grooming, 284
- all-optical edge grooming, 286
- all-optical grooming, 279
- all-optical wavelength grooming, 280
- amplifier, 124
- arbitrary alternate routing, 208

- bandwidth-blocking probability (BBP), 113
- bi-directional path networks, 69
- bi-directional ring networks, 69
- blocking probability, 193
 - end-to-end, 201
 - fiber route, 204, 206
 - g-link, 195, 197, 201, 204, 206, 208
 - path, 195, 197, 201, 208
 - segment, 204, 206
 - wavelink, 197, 199
 - active call, 207
 - passive call, 207
 - waveroute, 196, 197, 201, 204
- BXC, 254–256, 260, 261, 269, 271, 276

- call
 - active, 206
 - class, 198
 - multi-service, 198
 - passive, 206
 - service, 198
- carried load
 - fiber route, 205

- g-link, 200
 - active call, 208
 - passive call, 208
- path, 200
 - wavelink, 197, 201, 205
 - waveroute, 200, 205
- central office (CO), 107
- class- (n, k) call, 198
- complete lightpath cycle, 131
- complexity of Traffic Grooming, 57
- composite signal, 255
- connected lightpaths, 130
- control plane support
 - ASON, 21
 - mapping GMPLS, 26
 - call control, 31
 - hierarchical routing, 30
 - link resource manager, 28
 - mapping GMPLS, 27
 - model, 22, 26
 - resource discovery, 28
 - routing areas, 32
 - routing database, 28
 - subnetwork point, 26, 27
 - subnetwork point pool, 26, 27
 - UNI/NNI, 25
- GMPLS, 21
 - adaptation capability, 28
 - bi-directional LSP, 31
 - bundling, 28
 - call control, 31
 - exchange point, 25
 - explicit label, 31

- explicit route, 31
- forwarding adjacency, 28, 30
- generalized label, 27, 31
- hierarchical routing, 30
- hierarchy, 22
- label set, 31
- link management, 22, 29
- LSP, 22, 26
- mapping ASON, 26, 27
- MIB, 28
- model, 22, 26
- OPEX reduction, 21
- overlay/peer/augmented, 25
- resource discovery, 28
- routing protocols, 22
- RSVP-TE, 22
- signaling, 30
- signaling protocols, 22
- SRLG, 29
- support for VCAT/LCAS, 32
- switching capability, 27, 28
- TE link, 22, 26
- testbeds, 35
- upstream label, 31
- virtual link, 30
- grooming data plane
 - OTN, 22, 24
 - SONET/SDH, 22, 24
 - VCAT/LCAS, 24, 33, 34
- L1 VPN, 33
- PCE, 30, 34
- SLA/SLS, 33
- cost function, 261, 272
- cost per bit, 256
- crossbar switch, 17

- data centers, 253
- database backups, 253
- dedicated backup multiplexing, 152
- dedicated backup reservation, 150
- dedicated protection, 276
- demand time conflict reduction, 164
- dependable dynamic traffic grooming, 149
- deterministic, 254
- Digital Cross-Connect System (DCS), 15
- distribution
 - exponential, 111
 - Poisson, 111

- Dynamic Bandwidth Allocation (DBA), 51
- dynamic provisioning, 142
- dynamic restoration, 138
- dynamic traffic, 5

- electrical aggregation, 257
- electrical cross-connect, 254, 255
- electrical grooming, 254, 256, 257, 262, 272, 276
- elemental network topologies, 61
- End-to-end call loss probability, 197
- EXC, 254–257, 259, 260, 262, 266, 276

- Fiber Channel (FC), 108
- Fiber Connection (FICON), 108
- Fiber Cross-Connect (FXC), 12
- fixed-point equation, 197, 202

- g-link, 195, 196
 - carried load, 197
 - repeated, 209
 - support, 195
 - topology, 195
- general shared path protection, 144
- Generic Framing Procedure (GFP), 52, 108, 215
- grid networks, 69
- grooming, 254
- grooming and routing (GR), 123
- grooming lightpath, 257–259, 263, 264, 272, 273
- grooming link, 195
- grooming strategy, 256
- grooming switches, 15–19
- grooming, routing and wavelength assignment (GRWA), 124

- heuristics, 254
- hierarchical grooming
 - clustering, 79
 - general topology networks, 77
 - hub selection, 79
 - lightpath grooming, 83
 - logical topology, 81
 - performance results, 83
 - ring networks, 74
 - RWA, 82
 - star networks, 76

- torus networks, 76
- traffic routing, 81
- tree networks, 76
- hierarchical grooming algorithm, 79
- I/O ports, 256, 260
- Integer Linear Programming (ILP), 124
- inter-office (IOF), 107
- inverse multiplexing, 109
- Lagrangian relaxation, 124
- light tree, 18
- light-trail, 155, 165
 - dynamic, 168
 - static, 166
- lightpath, 2, 254–256, 260, 261, 264, 268–270, 272, 273
- lightpath cycle, 131
- lightpath sequence, 194
- Line Terminal Equipment (LTE), 59
- linear programming, 254
- Link Capacity Adjustment Scheme (LCAS), 51, 108
- link protection, 276
- link-based protection, 140
- logical topology, 2, 195
 - dynamic, 195
- management burden, 256
- Manhattan street networks, 69
- many-to-one traffic grooming, 227–232
 - exact approach, 229
 - heuristic approach, 231
- marginal demands, 259, 264
- matrix algebra, 263
- meta-heuristics, 254
- MG-OXC, 255, 260, 261
- mixed primary-backup grooming policy, 149
- monetary cost, 254
- Multi-Commodity Flow (MCF), 64
- multi-domain traffic grooming
 - conclusions, 251
 - introduction, 237
 - multi-domain standards, 238
 - open problems in multi-domain grooming, 245
 - research studies, 242
- multi-granularity switching, 255, 256
- multi-hop full grooming OXC, 17
- multi-hop partial grooming OXC, 18
- multi-layer, 255, 256, 262, 276
- Multi-Service Provisioning Platform (MSPP), 15
- multicast traffic grooming, 214–227
 - dynamic multicast traffic grooming, 222–227
 - enabling technology, 215
 - static multicast traffic grooming, 216–222
 - exact approaches, 216–220
 - heuristic approaches, 220–222
- multipoint traffic grooming, 213–233
 - types, 213
- network cost, 256, 261–263, 272, 276
- network design, 254
- network layer, 256, 261, 276
- node-edge incidence matrix, 126
- nominal capacity, 257
- NP*-Complete, 57, 112, 123
- objective function, 261
- OEO, 3
 - switching cost, 3, 60
- offered load
 - active call, 207
 - g-link, 200
 - passive call, 207
 - segment, 205
 - wavelink, 197, 201, 205
 - waveroute, 197
- optical burst switching (OBS), 287
- Optical Cross-Connect (OXC), 12, 114, 137
- optical grooming, 254, 256, 260–262, 268, 270, 273, 274, 276
- optical layer, 125
- optical node architectures, 9–19
- optical packet switching (OPS), 287
- optical switching, 13
- optical time-division multiplexing (OTDM), 284
 - principle, 285
- Optical Transport Network (OTN), 108
- optimisation techniques, 254
- opto-electronic, 256
- p*-Cycle protection, 141

- partial protection, 153
- partition, 269
- path protection, 276
- path restoration algorithm, 142
- path-based protection, 140
- performance model, 193
- permanent connection, 256
- permanent demands, 254
- physical layer, 125
- planning, 254
- Point-to-Point Protocol (PPP), 108
- power, 269
- pre-cross-connected shared path
 - protection, 145
- preplanned protection, 138
- protection design, 143
- provisioning, 254, 256, 273

- random traffic demands, 253
- random traffic grooming, 209
- random wavelength assignment, 209
- Reconfigurable OADM (ROADM), 280
 - multi-degree, 282
- reduced load approximation, 194, 196
- restricted sharing admission, 196, 198
- Routed Scheduled Band Group, 261, 270, 273, 274
- router, 255
- Routing and Wavelength Assignment (RWA), 3, 4

- scheduled demands, 253, 254, 276
- scheduled lightpath demands, 268
- scheduled services, 160
- SED, 254, 256, 257, 262–265, 267, 272, 273, 276
- segment, 204
- segregated primary-backup grooming
 - policy, 149
- self-routing fabric, 17
- service- k call, 198
- set-up date, 253, 254, 261, 262, 268
- shared backup multiplexing, 151
- shared backup reservation, 150
- shared bus, 16
- shared memory, 17
- shared protection, 276
- shelf, 269
- shortest path, 256, 262, 263, 268, 272

- single-hop grooming OXC, 17
- single-layer, 256
- SLD, 254, 256, 261, 262, 267–270, 273, 274, 276
- sliding scheduled traffic model, 162
- Soft Failure, 51
- solution space, 270
- SONET/SDH, 39
 - frame structures, *see* STS
 - layers, 40
 - Next Generation (NG-SONET/SDH), 47
 - rings
 - protection methods, 43
 - types, 43
 - standard rates, 41
 - Virtual Tributaries (VT), 43
- SONET/SDH ring grooming, 91–105
 - arbitrary traffic, 98–105
 - Bidirectional Line-Switched Ring (BLSR), 91
 - fixed traffic, 92–98
 - algorithms, 96–98
 - bounds, 93–96
 - hubbed rings, 95, 98
 - incremental traffic, 104
 - multi-ring topologies, 97
 - traffic types, 91
 - Unidirectional Path-Switched Ring (UPSR), 91
- source-edge incidence matrix, 126
- spider networks, 69
- star networks, 69
- state-of-the-art, 256
- static traffic grooming with min-max
 - objective, 71
- statistical multiplexing gain (SMG), 12
- STM, 41
- storage-area network, 108
- STS, 41
- subgraph-based routing, 140
- subgraph-based routing strategy, 142
- subwavelength demands, 3, 253–256
- survivability, 138
- switch ports, 254
- switching fabric, 255, 256
- switching granularity, 253

- tear-down date, 253, 254, 261, 262, 268, 276
- time diagram, 257, 272
- time-disjoint, 254, 276
- topology, 254, 257, 260–263
- traffic grooming, 159
 - dynamic, 193
 - multi-hop, 194
 - policy, 206
 - single-hop, 194
 - static, 193
- traffic pattern, 111
- transit traffic, 256
- transponder, 12, 124
- transponders, 254
- tree networks, 71
- tuple, 254, 261

- uncertainty, 254, 276
- unidirectional path networks, 62
- unidirectional ring networks, 69

- ventilation system, 269
- Virtual Concatenation (VC/VCAT), 48, 108, 109
- Virtual concatenation group (VCG), 48
- virtual topology, 2

- waveband, 253–256, 260, 261, 270, 271
- waveband cross-connect, 254, 255
- waveband switching, 261
- waveband-switching connection, 260, 261, 267, 268, 271
- wavelength, 255
- wavelength assignment, 260
- wavelength assignment (WA), 123
- wavelength capacity, 125
- wavelength continuity, 126
- wavelength continuity constraint, 260
- wavelength conversion, 204, 260
- wavelength cross-connect, 254, 255
- Wavelength Interchanging Cross-Connect (WIXC), 13
- Wavelength Selective Cross-Connect (WSXC), 13
- Wavelength-Band Cross-Connect (WBXC), 12
- wavelink, 195, 196, 198
 - repeated, 204, 209
- wavelink blocking model, 198
- waveroute, 195, 196, 198, 204
- WDM grooming, 193
- WXC, 254–256, 259–262, 266, 269, 271, 276