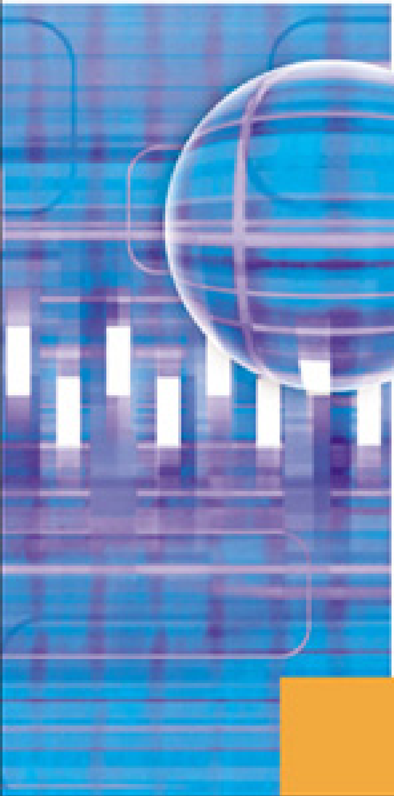
McGraw-Hill TELECOM
PROFESSIONAL

SONET-BASED METRO AREA NETWORKS

Planning and Designing the Next-Generation Provider Network



DANIEL MINOLI
PETER JOHNSON • EMMA MINOLI

SONET-Based Metro Area Networks

Planning and Designing the Next-Generation Provider Network

Daniel Minoli
Peter Johnson
Emma Minoli

McGraw-Hill
New York Chicago San Francisco Lisbon
London Madrid Mexico City Milan New Delhi
San Juan Seoul Singapore Sydney Toronto

McGraw-Hill

A Division of The ***McGraw-Hill*** *Companies*

Copyright © 2000 by The McGraw-Hill Companies. All rights reserved. Manufactured in the United States of America. Except as permitted under the United States Copyright Act of 1976, no part of this publication may be reproduced or distributed in any form or by any means, or stored in a database or retrieval system, without the prior written permission of the publisher.
0-07-140960-2

The material in this eBook also appears in the print version of this title: 0071402268

All trademarks are trademarks of their respective owners. Rather than put a trademark symbol after every occurrence of a trademarked name, we use names in an editorial fashion only, and to the benefit of the trademark owner, with no intention of infringement of the trademark. Where such designations appear in this book, they have been printed with initial caps.

McGraw-Hill eBooks are available at special quantity discounts to use as premiums and sales promotions, or for use in corporate training programs. For more information, please contact George Hoare, Special Sales, at george_hoare@mcgraw-hill.com or (212) 904-4069.

TERMS OF USE

This is a copyrighted work and The McGraw-Hill Companies, Inc. (“McGraw-Hill”) and its licensors reserve all rights in and to the work. Use of this work is subject to these terms. Except as permitted under the Copyright Act of 1976 and the right to store and retrieve one copy of the work, you may not decompile, disassemble, reverse engineer, reproduce, modify, create derivative works based upon, transmit, distribute, disseminate, sell, publish or sublicense the work or any part of it without McGraw-Hill’s prior consent. You may use the work for your own noncommercial and personal use; any other use of the work is strictly prohibited. Your right to use the work may be terminated if you fail to comply with these terms.

THE WORK IS PROVIDED “AS IS”. McGRAW-HILL AND ITS LICENSORS MAKE NO GUARANTEES OR WARRANTIES AS TO THE ACCURACY, ADEQUACY OR COMPLETENESS OF OR RESULTS TO BE OBTAINED FROM USING THE WORK, INCLUDING ANY INFORMATION THAT CAN BE ACCESSED THROUGH THE WORK VIA HYPERLINK OR OTHERWISE, AND EXPRESSLY DISCLAIM ANY WARRANTY, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. McGraw-Hill and its licensors do not warrant or guarantee that the functions contained in the work will meet your requirements or that its operation will be uninterrupted or error free. Neither McGraw-Hill nor its licensors shall be liable to you or anyone else for any inaccuracy, error or omission, regardless of cause, in the work or for any damages resulting therefrom. McGraw-Hill has no responsibility for the content of any information accessed through the work. Under no circumstances shall McGraw-Hill and/or its licensors be liable for any indirect, incidental, special, punitive, consequential or similar damages that result from the use of or inability to use the work, even if any of them has been advised of the possibility of such damages. This limitation of liability shall apply to any claim or cause whatsoever whether such claim or cause arises in contract, tort or otherwise.

DOI: 10.1036/0071409602

For Anna and her unwaivering help with InfoPort
(Daniel)

For Emmanuelle, my inspiration
(Peter)

For Peter, my inspiration
(Emma)

This page intentionally left blank.

CONTENTS

PREFACE

Industry observers have called 2001 a watershed year in the telecommunications and data communications industry. Five factors have relevance to a discourse on new-generation networks that planners may consider deploying in the near future: (1) The recent passing away of many legacy narrowband service providers such as competitive local exchange carriers (CLECs), digital subscriber line (DSL) local exchange carriers (DLECs), building local exchange carriers (BLECs), radio local exchange carriers (RLECs), Ethernet LECs (ELECs), and Internet service providers (ISPs), (2) the need for cost-effective but financially sound broadband service delivery, (3) the major recent advances in optical technologies, (4) the apparent glut of fiber facilities in the long-haul portion of the network but the absolute dearth of such facilities for broadband-services use in the local metro access/metro core segment of the network, and (5) the emergence of new access and backbone architectures, in particular Ethernet-based wide area network (WAN) approaches. According to these same observers, there will be a "changing of the guard in companies, architecture, products, technologies, and business models" going forward.

Major recent advances in optical technologies are poised to enable widescale deployment of broadband services. Optical networking and evolving Ethernet WAN technologies have the potential to foster a revolution in network design, services, and economics at least at the local metro access/metro core level. New service providers have emerged to apply these evolving technologies and position themselves to deliver new broadband services at the 100-, 1000-, and 10,000-Mbps range in the metro access/metro core portions of the network, as well as end to end, in a cost reduced fashion compared with services provided by the incumbents. We refer to them as metropolitan area service providers (MASPs) to focus on their role at the functional level.

A September 2001 front-page article in *NetworkWorld* magazine stated:

> . . . AT&T dives into metro Ethernet . . . AT&T will become the first of the big three wide area carriers to enter the Ethernet-based service . . . the announcement is significant because it represents the largest carrier embracing a potentially disruptive technology. End-to-end Ethernet could ultimately trash the cost conversions of traditional wide area datacom services, a fact that has caught the attention of many carriers. Metropolitan Ethernet services offer users an alternative to standard SONET OC-3c or OC-12c. There

> are fewer hassles handing off an Ethernet connection to a metropolitan or Ethernet network, and the cost is expected to be lower. There is every reason to believe that Ethernet services will become very popular . . . AT&T is not the first company to offer local Ethernet services . . . The key benefit of the service is that users hand us an Ethernet signal directly. They do not need SONET equipment or expertise . . . WorldCom has Ethernet service plans of its own . . .

These authors are publishing a companion set of two books that explore rapidly evolving architectures, products, technologies, and business models, particularly from a metro access/metro core perspective, where all the action now is. The two books will prove useful to planners in carrier, service provider, and enterprise environments. It will also prove useful for decision-makers, technology developers, and for students. It also should prove useful to Venture Capitalists who wish to deliver some profits to their investors.

Within the two books, one book looks at *traditional-but-fast-improving architectures* while the other looks at *evolving architectures*. Architectures, technologies, and services for high-end commercial applications are explored in detail in both books. This book looks at Synchronous Optical Network (SONET)/Synchronous Digital Hierarchy (SDH) and the powerful new Next-generation SONET, while the other book looks at Ethernet-based metropolitan area and wide area architecture such as Virtual LAN (VLAN), Rapid Spanning Tree Algorithm (RSTA), Resilient Packet Ring (RPR), wide-area 10 Gigabit Ethernet, and other Ethernet-based systems. The two books contrast these architectures to multilambda rings. Ring-versus-mesh-architectures are also studied.

This book, *SONET-Based Metro Area Networks: Planning and Designing the Provider Network*, focuses on the significant advances in optics and Next-generation SONET approaches, with features and services that can rival what the new Ethernet architectures can offer. Can Ethernet-based architecture scale in size, robustness, reliability, and carrier OAM&P capabilities as the Next-generation SONET systems?

Chapter 1 sets the stage for an assessment of the advances in SONET systems that enable them to compete with new architectures. Chapter 2 provides a short primer on optical technology. Chapter 3 looks at traditional broadband architectures, such as SONET/SDH, focusing on the self-healing robustness and OAM&P features. Chapter 4 looks at next-generation SONET and the advantages it affords; a number of new-generation carriers are going this route to provide advanced services. Chapter 5 focuses on optics and Dense Wavelength Division Multiplexing

(DWDM), a technology so far optimized for long-haul use. Chapter 6 looks at a number of other practical applicable technologies for Transparent LAN Service (TLS) delivery, such as Coarse WDM (CWDM). Chapter 7 covers evolving all-optical/intelligent optical networking technologies. Chapter 8 discusses wavelength services managed by Generalized Multiprotocol Label Switching (GMPLS). Chapter 9 discusses Free Space Optics, which some proponents are proposing as a metro access technology. Finally, Chapter 10 looks at practical design approaches and economics to validate the premise that these evolving technologies and architectures (both the Ethernet and the Next-generation SONET) can be reliably and sustainably deployed in the field. Some ELECs have already filed for Chapter 11 (and others will likely follow), indicating the absolute need for these providers and their CFOs to understand Economics 101. An acronym list is also provided to ensure that the reader has a built-in reference for the expanded names of many of the frequently used acronyms within the book.

Its companion, *Ethernet-Based Metro Area Networks: Planning and Designing the Provider Network* focuses on all the evolving Ethernet-based approaches. The book discusses the metro access/metro core environment, the market potential, and the required broadband services. Transparent LAN Service (TLS), in particular, is discussed along with a general view of technologies, applications, and opportunities. The book examines rearchitected metro rings and multiservice support, including Gigabit Ethernet (GbE). There is a close look at the performance of Ethernet-based systems. Virtual LANs (VLANs) are covered. Surveys of new data-oriented standards and their applications are provided, specifically, a survey of the Rapid Spanning Tree Protocol (RSTP). A short discussion on architectures with an eye to service availability is also included. The ELECs utilize GbE and VLAN techniques to deliver their services; hence, at this point in the book, a quick assessment of their approach is provided. Further chapters continue this discussion on even newer protocols by delving into the evolving Resilient Packet Ring (RPR) work, as well as examining 10 Gigabit Ethernet and exploring the question if carrier-class Ethernet services can emerge as a common bearer. A chapter also discusses the Fibre Channel Standard and its support over WANs. The book also looks at a brand-new initiative to expand Ethernet to residential and SOHO applications; this is known as the Ethernet in the First Mile (EFM) initiative. An acronym list is also provided to ensure that the reader has a built-in reference for the expanded names of many of the frequently used acronyms within the book.

Note that we have chosen to avoid an up-front discussion on the myriad of academic architectures that are or have been advanced, and pursue long taxonomies; we cover the issue of architecture within the context of each technology, and only focus on those architectures that show real near-term commercial possibilities.

Also note that we have drawn information and concepts from various industry sources, which we acknowledge and thank at this juncture. Our value, we hope, is to assimilate, evaluate, highlight, pare down, compare, order by importance, and classify these technologies, alternatives, and proposals in terms of real utility in building a network for a new-generation telco that brings positive net income to the bottom line over a short business cycle.

ACKNOWLEDGMENTS

The authors would like to thank Data Connection Ltd., and specifically thank white paper authors Neil Jerram and Adrian Farrel of Data Connection Ltd., for materials supplied for sections of Chapter 7, “All-Optical Networks,” and Chapter 8, “GMPLS in Optical Networks.”

ABOUT THE AUTHORS

Daniel Minoli Mr. Minoli, Managing Director, *Econsultex Inc.*, has 27 years of marquee corporate experience in technology and applications related to, and focused on, broadband enterprise and carrier communications, Metro Ethernet and next-gen optics, LAN technology, voice-over-IP, Internet and intranet architectures and design, wireless LANs and WANs, e-commerce, and multimedia applications. Most recently, Mr. Minoli was founder and/or co-founder of several high-tech companies including *InfoPort Communications Group*, an Optical & Gigabit Ethernet Metro Carrier. In addition to a full-time corporate career, Mr. Minoli has taught graduate Optical Networking/datacom, Voice-over-IP, multimedia, and e-commerce at Stevens Institute, NYU, CMU, and Monmouth University as an adjunct professor for 18 years. He has been a technical analyst for Gartner Dataquest (Datapro) for 17 years. He has also been a consultant for numerous Venture Capitalists' high-tech investments (exceeding ≈$125M of investment). Mr. Minoli enjoys over 3,000 Web hits with his name (such as, Google.com "Minoli Daniel"). He has been a columnist for trade periodicals (*ComputerWorld*, *NetworkWorld*, *Network Computing*), has written dozens of technology reports for *Datapro Corporation* and the *Gartner Group*, and has written five full-fledged market reports for *Probe Research Corporation*. Mr. Minoli has spoken at over 75 conferences and published over 300 articles. He is often sought out for advice by companies, patent attorneys, and Venture Capitalists, and has been involved in mezzanine investments, providing in-depth reviews of technology and market baseline for high-tech companies in the multimedia, digital video, CTI/Java, VSAT, and telemedicine arenas.

In the 1990s he was a senior executive (VP Packet Services) at *TCG/AT&T*, where he deployed several large data networks with cumulative P&L $250M capex, $75M opex, $125M direct revenue, and $4B of revenue impacted. Mr. Minoli started the Broadband/IP Services operation at Teleport Communications Group in late 1994; he was "Broadband Data Employee #2" and built the operation to a $25M/yr business (the unit value was $250 to 300M). Mr. Minoli's team deployed 2,000 backbone/concentration/access routers and 100 ATM/Frame relay switches in 20 cities in 5 years; the team turned up 1,500 active broadband ports and secured 400 broadband customers. Prior to AT&T/TCG, Mr. Minoli worked at DVI Communications (Principal Consultant), where he managed the deployment of three major corporate ATM networks, including one for the Federal Reserve Bank. From 1985 to 1994 Mr. Minoli worked at Bellcore/Telcordia (Technology Manager) where he thoroughly researched all aspects of broadband data networking, from service creation to standards writing; from architecture design to marketing; and so on. In the mid-1980s he worked at Prudential Securities (Assistant VP), deploying dozens of data networks, including a 300-node VSAT branch data network. At ITT Worldcom (1980–84) he deployed a fax-over-packet network, and automated nearly all carrier operations by developing over a dozen OSSs, including Customer Care/CRM. At Bell Labs he worked on internal and public packet networks, and at Network Analysis Corporation he undertook ARPA work on voice-over-packet, wireless networks, and integrated communications.

Peter Johnson Mr. Johnson is a telecommunications business and technology professional. He has worked in a multitude of diverse positions within the industry, including those of independent consultant, business analyst, senior data engineer, lead project manager, senior product manager, and senior marketing manager. He has run nationwide projects involving over 200 locations, managed entire departments, helped launch new companies, and provided strategic research and analysis to help companies obtain additional funding. Mr. Johnson's experience in telecommunications has spanned across many companies in several different areas, including those of ILEC, RBOC, IXC, CLEC, start-up optical carrier, start-up optical equipment vendor, and multiple business and technology consulting firms. Mr. Johnson also currently holds a position on the board of directors of a small business and technology consulting company based in the Northeast.

Emma Minoli Ms. Emma Minoli is CEO and Founder of *Red Hill Consulting*, a high-tech consultancy specializing in the area of Voice Over IP, multimedia, and e-commerce technologies. *Red Hill Consulting* also focuses on marketing and PR/advocacy for these technologies. During the past 6 years clients have included companies such as AT&T, J. D. Edwards, Chase Manhattan Bank, InfoPort Communications Group, and EtherCarrier Communications. Ms. Minoli has co-authored several books including *The Telecommunications Handbook*—Digital Video section (CRC Press/IEEE Press), 1999; *Delivering Voice over IP Networks* (John Wiley & Sons), 1998; *Delivering Voice over Frame Relay and ATM* (John Wiley & Sons), 1998; and *The Web Commerce Technology Handbook* (McGraw-Hill), 1997.

CHAPTER 1

Advances and Opportunities in Next-Generation SONET and Other Optical Architectures

In spite of some temporary slowdowns at the beginning of this new decade end-user organizations continue to have increasing bandwidth needs for intranets, connectivity to backup and storage sites, access to the Internet, and multimedia applications. As these organizations deploy local area networks (LANs) comprising switched segments based on 100-, 1,000- or 10,000-Mbps technology, there is a correlated need for broadband wide area network (WAN) services, metropolitan area network (MAN) services, and access network services.[1] As a rule of thumb, the WAN/access requirement is equal to the aggregate LAN speed divided by 100.[2] Hence, a corporate site with an aggregate LAN traffic of 1 Gbps (for example, 100 users on 10-Mbps LAN segments) will need 10 Mbps of sustained WAN throughput. In a few years, when these 100 users are on 100-Mbps segments, the corresponding WAN requirement from such a site will be 100 Mbps. Network survivability, high availability, security, and fast restoration are also key desiderata of end users regardless of their bandwidth needs. Some people see wavelength leasing and wavelength on demand as evolving user requirements as well. These services, however, fall more in the circuit-switched mode of operation when contrasted with the newly introduced packet-mode services based on Ethernet and Ethernet-like technologies that are receiving considerable industry attention.

Users hope to be the beneficiaries of the bandwidth glut that was being reported in the press in the 2001–2002 time period. The need for reliable broadband services is obvious. At least two parochial uses of the term "broadband" have evolved in the recent past. Some people have applied broadband to mean xDSL Internet access at what turned out to be 384, 768, or at most 1,544 Kbps. This is far from being broadband in any sense of the word. Others have applied broadband to mean cable-TV-based residential services on digital cable. The proper use of the word "broadband" (according to the International Telecommunications Union [ITU]) is for a service that has a speed greater than a T1 (1.544 Mbps); hence, to qualify as broadband, a service needs at least 3.088 Mbps (T1C), 6.312 Mbps (T2), 10 Mbps, or more. In fact, today many people consider 100 Mbps, 1 Gbps, or 10 Gbps to be the reference speeds for broadband. Current technology even supports speeds of 40 to 320 Gbps; these, however, are more backbone network speeds than user-level speeds.

The purpose of this book is to survey the architectures and technologies needed to deliver reliable, cost-effective, sustainable, secure, and carrier-grade broadband services to commercial customers. There are both *traditional-but-fast-advancing* solutions and *new-but-not-yet-large-scale-tested* solutions. This book is part of a two-volume set on this topic. This book examines traditional-but-fast-advancing architectures and technologies; the companion book[3] examines the new-but-not-yet-large-scale-tested solutions, which promise to alter the landscape of the metro access/metro core environment. There is virtually no overlap between the two books, except for the few introductory concepts in the following subsection.

Currently, service providers, enterprise network planners, and technology developers are being forced to rethink transport, routing, switching, and traffic management for the networks now in the process of being planned and/or deployed. Changes in architectures and infrastructures for broadband deployment could be far reaching. These implications are discussed in the material that follows.

1.1 Background

Industry observers have called 2001 a watershed year in the telecommunications and data communications industry. Five factors have relevance to a discourse on the new-generation networks that planners may consider deploying in the near future: (1) the recent passing away of many legacy narrowband service providers such as Competitive Local Exchange Carriers (CLECs), Digital Subscriber Line LECs (DLECs), Building LECs (BLECs), Radio LECs (RLECs), Ethernet LECs (ELECs), and Internet service providers (ISPs); (2) the need for cost-effective, but financially sound broadband service delivery; (3) the major recent advances in optical technologies; (4) the apparent glut of fiber facilities in the long-haul portion of the network, but the absolute dearth of such facilities for the use of broadband services in the local metro access/metro core segment of the network; and (5) the emergence of new access and backbone architectures, in particular Ethernet-based WAN approaches. A sixth factor has also become pertinent of late: the need to build distributed, disaster-proof broadband communication networks that enable users to have backup sites in low-rise suburban environments and the related requirement to bring the underlying fiber infrastructures to these suburban buildings.

As end-user organizations deploy LANs that are comprised of Gigabit Ethernet (GbE) and 10 GbE technology, there is a correlated need for broadband network services, as depicted in Figure 1-1. Figure 1-2 shows an example of a demand forecast. According to this figure, there will be 70 million GbE ports in 2004 worldwide. We estimate that 55 percent of the deployment will be in the United States, or 38.5 million ports. The average office building supports 100 workers. This implies that there would be approximately 385,000 commercial office buildings in the United States needing support for broadband services by 2004 (there are approximately 735,000 office buildings in the United States as of 2000). According to Nortel, optical Ethernet services in the MAN are expected to take an

Figure 1-1 Current bottlenecks in metro access/metro core networks.

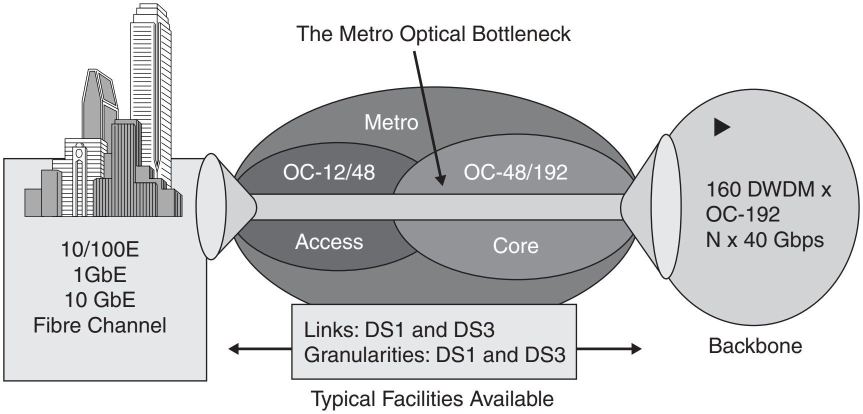
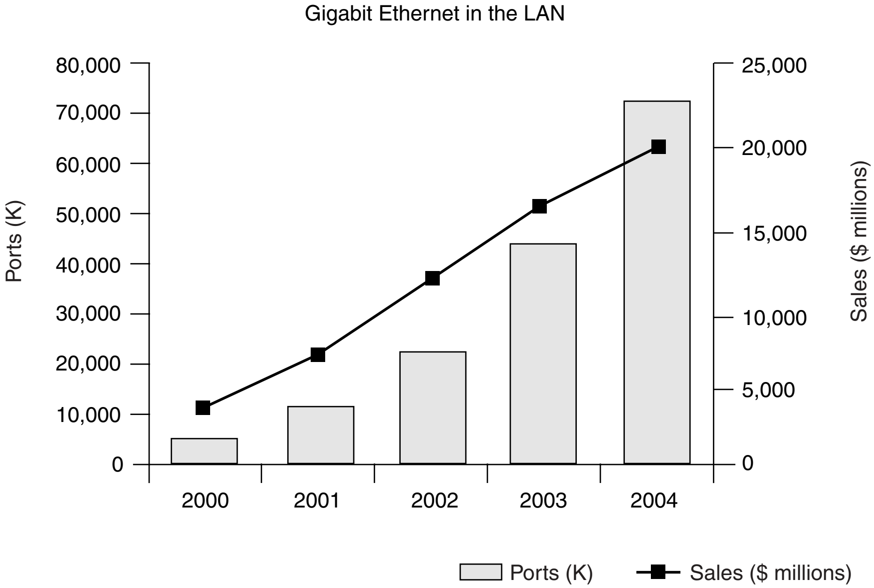


Figure 1-2 Drivers for broadband requirements in the MAN/WAN. Source: Cahners In-Stat Group, Sept 2000.



increasing share of the lit services market, rising to 20 percent by 2005, representing approximately $6 billion in revenues.[4] Even if this prediction about Ethernet in the MAN came true, it would still leave 80 percent of the market, or a market of $24 billion, to the traditional technologies.

Through major recent advances, optical technologies are poised to enable the wide-scale deployment of broadband services when the proper end-to-end focus and architectures are taken into account as part of the design process. Several technologies have emerged for providers who want to sell new metro access services to businesses. New metro access services can use Ethernet, dense wave division multiplexing (DWDM), passive optical network (PON), and data-aware Synchronous Optical Network/Synchronous Digital Hierarchy (SONET/SDH) (next-generation SONET) technologies to connect businesses to optical transport for WAN and Internet access. Residential customers will ultimately have access to fiber and intelligent optical-networking gear, but in the near future, customers will have to use DSL, cable, and fixed wireless access technologies to meet their demand for increased bandwidth.[5]

Optical networking and evolving Ethernet WAN technologies have the potential, at least in principle, to foster changes in network design, services, and economics. New service providers have emerged that deploy both the evolving and the new technologies, and position themselves to deliver new broadband services at the 100-, 1,000-, and 10,000-Mbps range in the metro access/metro core portions of the network, as well as end-to-end over the long haul. We also refer to them as metropolitan area service providers (MASPs) to focus on their role at the functional level. Some MASPs, particularly the ones that have labeled themselves ELECs, may or may not have a future, due to their limited service scope and restrictive architecture. Carriers with a solid business plan do have a future. Incumbent providers are also looking into new technologies, although they have a major embedded base of previous-generation systems supporting narrowband applications.[6] Figure 1-3 encapsulates some of these needs.[7]

Figure 1-3
Challenges of today's networks.

- **End Customer Requirements:**
  - – Increasingly faster and transparent networks
  - – Reliability and 100% uptime
  - – Low-cost bandwidth only when they need it
  - – Low-cost network services

- **Carrier Requirements:**
  - – Higher channel counts → >80 λs, 25-50 GHz
  - – Higher data rates → 10–40 Gbps in LH and ULH; 10 Gbps in metro
  - – Longer reach → >3500km (ULH)
  - – Less 'dumb,' static networks → Tunable networks
  - – Lower fixed cost of deployment, and decreased complexity in planning and management of networks
  - – Establish right-sized, net-income producing metropolitan networks optimized for broadband data applications

LH = Long Haul
ULH = Ultra Long Haul

The emerging optical-networking systems and services have the potential to enable metro carriers to scale up their infrastructure to meet the expansive demand at the metro-access-to-metro-core confluence point. This is apropos because there is currently a severe bottleneck in the metropolitan access networks that serve as the bridges between the long-distance core networks (where reportedly there is a bandwidth glut) and the end-user building and campus networks. Any hope of truly delivering broadband services to paying customers resides in addressing this issue, not just on pure technology considerations or an abstract view of networks without a proper end-to-end view.

The foremost, absolutely necessary, and mandatory considerations in delivering reliable, ubiquitous, and sustainable broadband services are the following:

1. The problem has to be conceived as an end-to-end challenge—just solving a portion of the problem is to no avail. Broadband connectivity must include the building riser fiber, the building access fiber, the metro access subnetwork, and the metro core backbone. For example, just distributing food staples to the Bronx Wholesale Market (or other major markets associated with other cities) and not finding a way to get the food to every store, restaurant, and home where it is needed is not achieving the intended goal. Just bringing hydroelectric power to Westchester, New York, and not finding a way to get the electrical power to every building, home, and factory in New York City where it is needed is not achieving the intended goal. Just bringing water to a city reservoir without bringing the water to every neighborhood, home, school, restaurant, factory, and hospital where it is needed is not achieving the intended goal. Therefore, just having broadband DWDM-based networks in the long haul, metro core, or point-of-presence-to-point-of-presence (POP-to-POP) segment without bringing the fiber connectivity to major commercial neighborhoods, office buildings, hospitals, industrial campuses, or college buildings is not achieving the intended goal.
2. The equipment cost is typically 15 percent of the total cost to deliver a service. Hence, a 20 percent decrease in equipment cost only results in a 3 percent decrease in total service cost. A new technology and/or a new piece of equipment needs to show deeper savings to really impact the bottom line. For example, it should reduce equipment costs by, say, 100 percent and/or reduce operations cost by 100 percent (which is typically 20 percent of the cost to deliver service). In this case, the bottom-line savings would be 17.5 percent, which might be enough for a customer to consider a switch in service/service provider.[8]

3. Reliability, security, and carrier-class-grade features are a must. Campus-derived service offerings by ELECs are arguably not sufficiently robust or secure. As users pack larger portions of their corporate traffic onto fewer broadband facilities, high levels of reliability and security become imperative.
4. Metro broadband communication is neither free nor cheap. End users requiring broadband at the local level need to comprehend the cost issues. With the lavish[9] long-haul investments that have appeared in the recent past, the price of long-haul services has gone down "as much as 92 percent . . . with fire-sale price offerings by less solvent competitors."[10] Unfortunately, these reductions in prices in long-haul services coupled with free-reign hype on the part of startup carriers seeking venture capitalist funds foster an unattainable expectation in end users for very low prices in metro access/metro core. Low prices have not been achieved yet in the metro access/metro core space partially for intellectual reasons. The focus of architects, designers, technologists, and venture capitalists has not been where it needs to be. The focus has been on developing an array of new Layer 2 protocols of all kinds rather than on lowering the per-building fiber access cost, which is comprised of the spur cost and the in-building equipment cost. However, there are also endogenous issues. Local services (such as a cab ride at $1.50 per mile) are intrinsically more expensive than long-haul services (such as air transportation, which is generally $0.25 per mile) because of the much smaller base upon which facility costs can be dispersed (for example, an access fiber can be shared among a handful of users—say, one to four users—whereas a long-haul fiber can be shared among several thousand users) and because of the economy of scale of the equipment that has to be used. Finally, there are regulatory and open-competition reasons.

Industry efforts need to focus on bringing innovative technology to the metro edge of the public service network, which as noted where nearly all the bandwidth scarcity and connectivity bottlenecks currently are. These optical systems need to deliver an optimized blend of user-preferred Ethernet access protocols, broadband speeds, cost-effective fiber-optic pair gain, multimedia support, and distance bridging via a family of price-performance-specific field-upgradeable platforms, enabling next-generation carriers, incumbents, and ISPs to efficiently deploy end-to-end Ethernet services (as well as traditional private-line services) faster, farther, and more flexibly than possible at the present time. However, although new Ethernet-based technologies are being applied to the MAN environment, traditional architectures and technologies, specifically SONET/SDH,[11] are making major strides in improving their price-performance points, sustaining the assault from the new alternatives, and cementing themselves as formidable embedded solutions. The SONET/SDH transmission protocol is used by carriers seeking to provide carrier-class services because it provides the needed Operations, Administration, Maintenance, and Provisioning (OAM&P) capabilities. Such capabilities reduce the operations cost, which nearly always exceeds the amortized equipment cost, particularly for large-deployment environments. Some of the simplistic management tools available on campus-derived architectures (for example, virtual LAN [VLAN] architectures) become significantly problematic from a productivity standpoint when the base of network elements reaches into the hundreds range. Both the topological and the OAM&P capabilities of SONET are needed by service providers to support large-scale networks.

Users and providers consider Ethernet to be a well-known, well-understood technology with a high comfort level among technical experts and network operators. Before deciding whether to use Ethernet-backbone-based networks or next-generation SONET-based networks to deliver new broadband services, service providers should first understand the issues and opportunities surrounding end-to-end Ethernet networks and/or transparent LAN service (TLS),[12] and then begin to incorporate Ethernet-based paradigms into their service development and deployment plans. Only then can they ascertain the extent of the impact on the proposed network architecture. And more than ever, carriers need to be mindful of their network design actions in regards to profitability.

A general data network infrastructure usually has three different levels:

1. *Access networks* include the local access facilities between a service-providing node (SN), central office (CO), or POP and the customer premises network. The customer premises network is comprised of a building, riser, and campus elements. Ethernet is the dominant technology for the LAN at both the floor and backbone levels. The access link is the last/first mile between the metro core/WAN network and the customer premises. Today, the provider usually owns the access network. The access link is mostly based on narrowband copper transmission technologies. The speed of the access link is therefore lower than Ethernet, Fast Ethernet, GbE, and 10 GbE speeds. Fiber deployment in major metro areas and advances in last-mile access technologies, such as wireless broadband and free-space optics (FSO), aim at removing the access network bottleneck. Unfortunately, today only a very small fraction (a few percentage points) of buildings are

actually on fiber links to the metro core network, and alternative access technologies are not proven or widely deployed.

2. *Metro core networks* interface with both the local metro access network and the WAN network. Historically, SONET-based transmission systems have provided MAN connectivity. These systems were originally designed and built to carry voice traffic. As such, they operate in a time-division multiplexing (TDM) mode at the physical (PHY) layer and rely on an added data link layer (DLL) protocol such as the Point-to-Point Protocol (PPP), Asynchronous Transfer Mode (ATM), or Frame Relay. Most of the traffic growth in these networks is now due to data applications. The critical question is whether the embedded SONET transport infrastructure (which represents billions of dollars of partially depreciated assets in an estimated 100,000 rings in the United States)[13] is adequate and/or optimized for data traffic. For a reference point, a typical GbE-based metro backbone deployed by the ELECs uses a trivialized statistical TDM (STDM) mode at the PHY layer[14] and a Media Access Control (MAC) protocol at the DLL level. Metro service providers are looking for cost-efficient, data-optimized solutions. The question then becomes whether a TDM PHY plus a PPP (or ATM) DLL is better than a trivialized STDM PHY plus a MAC DLL, or vice versa. The key to the answer is simple: Follow the money. In rendering an end-to-end reliable secure monitored service, where are the cost components that comprise the total cost? Is the cost in the protocol engine or is it in the physical fiber link from the building to the metro core network?
3. *WANs* have been implemented using a variety of architectures and technologies, but recently most have been implemented by deploying DWDM systems and securing OC-*x* links (for example, OC-3s, OC-12s, and OC-48s) to overlay data networks onto the bearer transport network.

In MANs and WANs, there is a nontrivial deployment of telco-managed fiber-optic rings for narrowband services. These rings are currently using telephony-centric protocols (such as TR-303) that do not scale to the demands of packet networks in general and broadband in particular. These demands include reducing equipment and operational cost, improving the speed of deployment, supporting bandwidth allocation and throughput, and providing resiliency to faults. Emerging optical products must be optimized for the metro-edge portion of the network, where commercial buildings (typically in the range of several thousand per city, or about 735,000 nationwide, as of 2000) need to be connected in an effective and high-speed manner to the carrier's SNs.

There has been a lot of hype in the recent past about metro DWDM. When the long-haul industry saw major retrenchments at the turn of the decade, optical vendors took the easy course of taking the equipment that they had developed for long-haul applications and pasting a metro DWDM label onto the equipment. This was done while generating new marketing collaterals rather than redeveloping equipment that is optimized for metro access/metro core applications (as would have been appropriate). We have argued that the window of opportunity of metro DWDM is very limited. Finally, the following quotes bring some satisfaction: "With so much unused fiber, when will metro nets really need WDM? What are the requirements for 40-Gbps systems? What are the new economic trade-offs between transparency and grooming? And which of the new service providers will succeed?"[10] Some brandname vendors have products that would cost several hundred thousand dollars per building. This is totally untenable because carriers generally need to keep their in-building costs at no more than $20,000 to have any chance of being profitable. Furthermore, the high building cost is multiplicative—every building in the ring would require the expensive optics. Although high-cost DWDM might be acceptable for point-to-point applications (not rings with 8, 16, or 32 buildings on them) and for long haul, it is completely unacceptable in the metro space.

Observers cite the following drivers for the introduction of new metro optical architectures and technologies:

- End-user drivers/issues
    - Increased capacity demand at the access level
    - Demand for lower-cost, higher-reliability, and diversified facilities
    - Fiber-based services available in more locations
    - Decreased competition (fewer providers around)
- Service provider factors
    - Need for decreased capex and opex, and requirement for improved bottom line
    - Revenue opportunities in optical layer
    - Network simplification and equipment consolidation
    - Mesh network efficiency and flexibility
    - New optical-networking technologies continually entering the market
    - Shorter end-user service contracts, which implies the need for faster depreciation

Table 1-1 depicts the typical end-user applications in terms of a number of key service parameters. The various metro architectures need to support

Table 1-1

Typical End-User Applications

	Data	Voice	Video
Capacity	Narrowband/ traditional apps Broadband/new apps	Narrowband	Broadband
Bursty	Yes	No	No
Tolerance for delay or jitter	Medium	Low	Low
Quality of service (QoS)	For some applications	Deterministic always	Deterministic always
Restoration requirements	Medium	High	Medium
Fault tolerance	Medium	Low	Low/medium

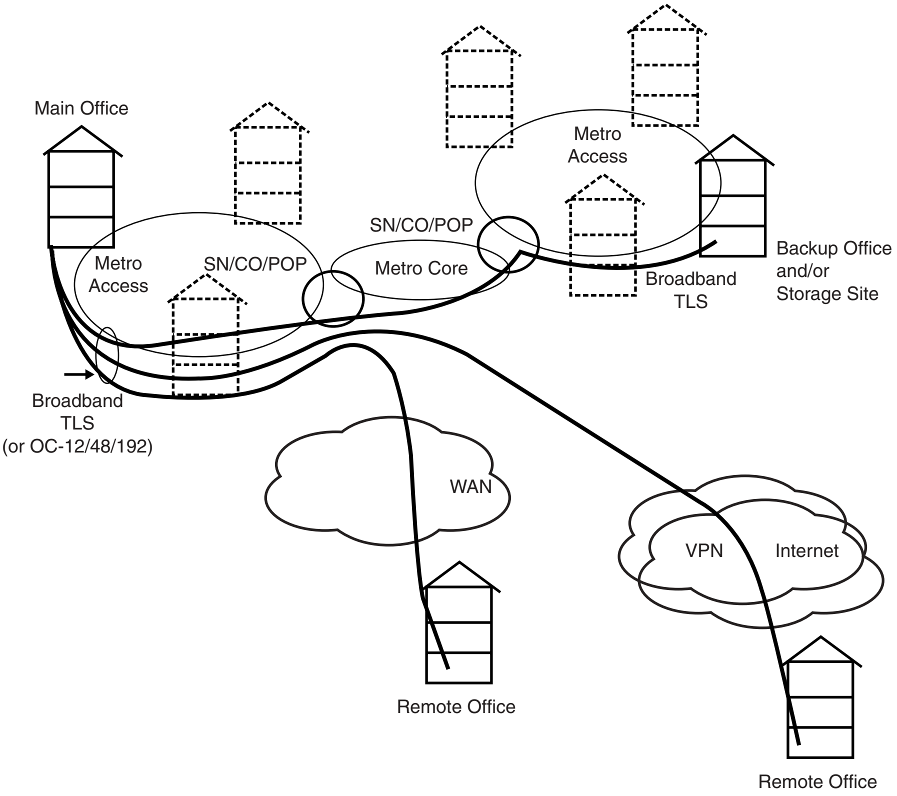
some or possibly all of these requirements. Broadband user-level applications include intranet data networking, storage and backup site connectivity, high-speed Internet access, and multimedia applications (see Figure 1-4 for examples).

In spite of recent advances in optical technologies, an intrinsic need to disruptively increase the available metro/WAN throughput remains. A trend is developing in the move toward full-speed, end-to-end 1 and 10 GbE in the WAN to support native Ethernet and IP over optics communication networks that next-generation carriers will set out to deploy in the next couple of years. This raises a pertinent question: Does the need and/or desire to deliver an Ethernet handoff to the customer require an Ethernet at the core metro infrastructure, or can such handoffs be supported more cost-effectively and reliably using a next-generation SONET platform that has Ethernet only at the edge?

Up to 70 percent of corporate WAN traffic is LAN data and less than 10 percent of WAN traffic is voice. Yet up to 70 percent of WAN links use DS1 and DS3 private-line, voice-oriented circuits to link up to other metro corporate sites or to link into the backbone.[15] Hence, the challenge for carriers is how to migrate from their current TDM/ATM-based networks to a world dominated by packet traffic driven by the growth of the Internet (which according to the best data is still growing at 80 percent a year although there may be some slowdowns in the future).[16]

As part of an assessment of the kind of architectures that carriers should deploy for supporting evolving networks, it is interesting to note some research at the time of this writing. Communications Industry Researchers,

Figure 1-4
Broadband/TLS applications.

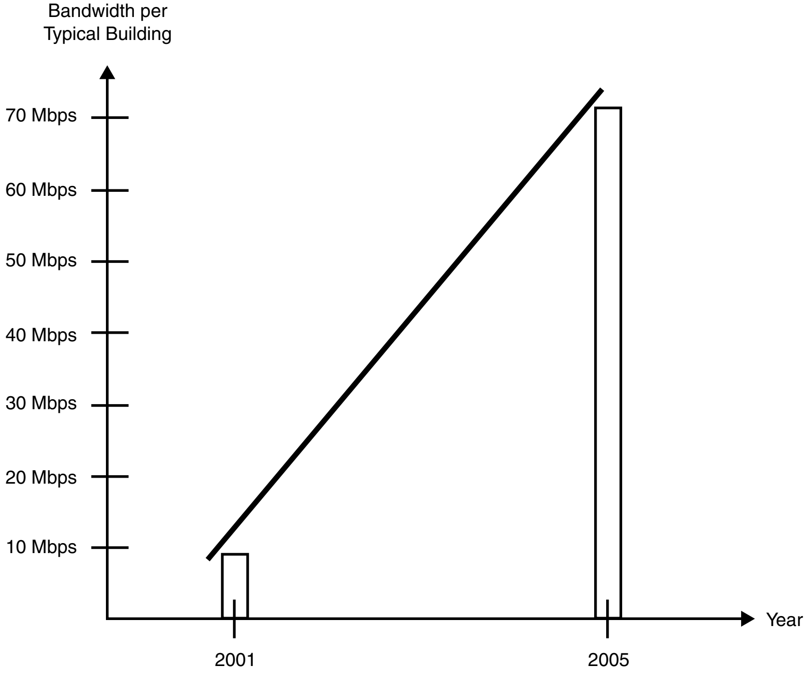


Inc. (CIR), an industry analysis firm, released a report[17] on the optical-networking strategies of two U.S. Incumbent Local Exchange Carriers (ILECs). Included in the report is an analysis of the wholesale wavelength services market. CIR said that it "thought the wholesale lambda services market would remain limited to the CLEC space for the foreseeable future. Some of the CLECs, such as metro carrier's carriers like Sigma and Telseon, are seeing more interest in wholesale lambda services at OC-48 rates and higher than they are in their retail Ethernet businesses. And, while both Yipes and Cogent like to promote themselves as retailers, major portions of their revenue is wholesale business." This observation points out that traditional SONET-based handoffs may be more prevalent in the near future than Ethernet handoffs. The fate of some of the ELECs may in fact reinforce this hypothesis.

Many of the current traditionally built networks deliver a data service known as *private line*. Private-line services are billable, physically (TDM) separated point-to-point circuits that are directly provisioned into segregated bandwidths in the intermachine trunks on the transmission infrastructure: DS1/DS3. Virtual private-line services are billable, virtually separated point-to-point circuits that are carried over (encapsulated or tunneled) common shared bandwidths provisioned in the intermachine trunks on the transmission infrastructure; examples include Frame Relay, ATM, and VLAN GbE.

Despite the weak telecommunications market and the slow economy of the early 2000s, the new metro access market relatively remained a hot spot at this writing for incumbent carriers, startup providers, and equipment manufacturers looking to create value for their customers. There is reportedly a large market potential for broadband services. Fiber to the Home (FTTH) systems are expected to grow at a decent pace during the decade. The number of MAN/WAN broadband (greater than 10 Mbps) terminations is expected to exceed 2 million by 2005.[18] Some people forecast 5 million FTTH subscribers in 2005 and 22 million in 2009.[19] Others quote 2.6 million FTTH subscribers with an additional 1.9 million subscribers with Fiber to the Curb (FTTC).[20] The average data rate per building is expected to grow from 10 Mbps (2001) to greater than 250 Mbps (2005) according to some, and from 10 Mbps to 73 Mbps per average customer according to the authors of this book (see Figure 1-5); others forecast even

Figure 1-5 Bandwidth requirements per building as a function of time.



higher numbers. According to published sources such as the Yankee Group, the market for wavelength services is expected to reach $1 billion by 2004.[21]

On the equipment front, in the second quarter of 2001 worldwide revenues for optical last-mile Customer Premises Equipment (CPE) reached $75 million, and they are expected to grow to $138 million by the second quarter of 2002, with Ethernet CPEs accounting for 18 and 37 percent of these revenues, respectively.[5] According to CIR, the market opportunity for networking vendors supplying SONET, DWDM, and optical Ethernet for U.S. metro access/metro core applications could reach $2.5 billion by 2005.[22] Market forecasts are typically optimistic; we have quoted herewith some of the best possible sources. In other sections of the book, we occasionally provide smaller forecasts based on our own assessment of the potential opportunities.

It is a well-known fact that less than 5 percent of commercial buildings in the United States with more than 5,000 square feet of space have fiber connectivity (only between 21,000 and 35,000, depending on the estimate, out of the 735,000 commercial office buildings in the United States, not to mention the entire set of 4.6 million commercial buildings).[23,24] Therefore, there is a tremendous opportunity in the next 5 to 10 years to develop products that will support the eventual mandatory connectivity that all of these buildings will need to support their existence in an information/Internet-based economy. The fiber glut of the early 2000s only existed in the long-haul environment, if at all, not in metropolitan network environments.

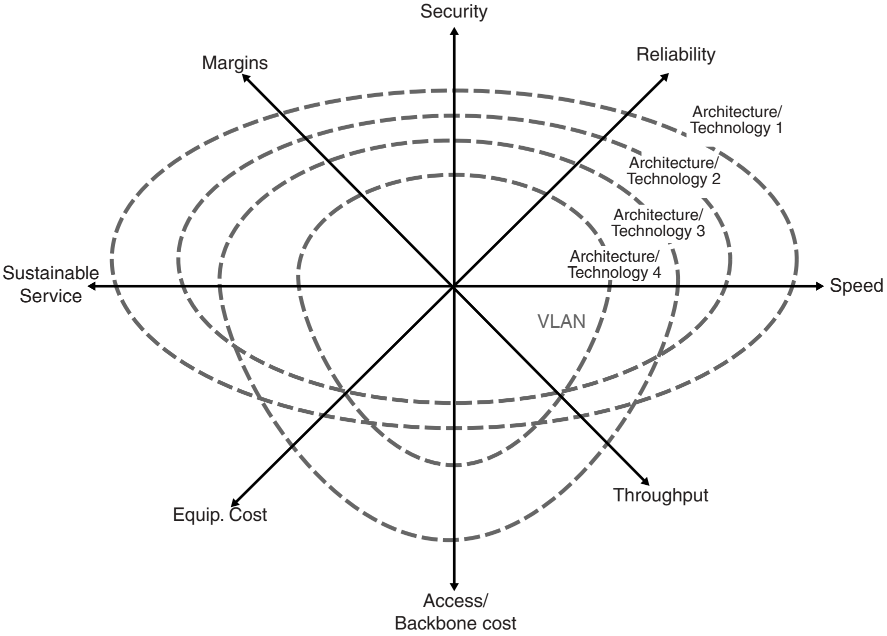
Without any doubt, broadband penetration will grow for the rest of the decade. It is worth noting, however, that a debate was under way at the time of this writing regarding the growth of the Internet. One source summarized the debate as follows: " . . . we have learned that while there is little evidence to support the 'urban legend' of an Internet that doubles every 3–4 months, neither we find any evidence that the Internet has engaged in a 'relentless process of slowing' as alleged by a prestigious Wall St. firm. And the reason we can't find any evidence of the latter is because there isn't any."[25] Based on what is described as the "first scientific study of Internet growth since 1996," Caspian Networks reported a 2001 growth at an annual rate of 4 times a year through early 2001, equating to a 200 percent yearly growth[26] (others have quoted a growth of 100 percent per year). (The doubling of the rate every 3 months would equate to an annual growth of 700 percent per year.) However, some carriers did report a slowdown to a 120 percent growth a year, while the financial press offers lower figures particularly for 2002.[25] Caspian Networks estimates that 80 percent of the IP traffic is enterprise driven from corporations that are increasingly turning to the Internet for cost reductions. The Yankee Group estimated that 21 percent of North American multinational corporations (for example, Global 500) had an OC-3 or higher network connection in 2001; that number is projected to reach 34 percent by 2003. Although there was overbuilding in 2000, proponents see renewed growth beyond 2002, making statements such as "lit fibers are heavily loaded . . . network performance will degrade. Any carrier that fails to upgrade in time risks losing market share and revenue to competitors."[25]

Optical-networking challenges are increasingly geared not just at obtaining higher rates,[27] but also at provisioning services effectively and reliably to end users. This is accomplished with a full set of OAM&P capabilities in the architecture and/or equipment. The end users and their applications fuel the demand for bandwidth. The rates currently found in high-end enterprise networks, storage area networks (SANs), application service providers (ASPs), and web-hosting sites are in the tens of gigabits per second, rivaling the rates carried in core systems and prompting the use of optical networking in such high-rate LANs. Simultaneously, applications such as site and/or server mirroring and corporate intranets require these high rates to be supported over new or existing MANs.[11]

1.2 The Panoply of Technologies and Architectures

MANs play a vital role in connecting users to WAN services. Traditionally, MANs have been circuit-switched networks that are optimized for voice traffic using the SONET/SDH circuit-switched layer. LANs, on the other hand, are packet-switched networks that are optimized for data traffic using Ethernet technology. This is quickly becoming the dominant traffic source driving bandwidth demands in MANs and WANs. The consequence of carrying packet traffic on circuit-switched networks is that it is inefficient at the bandwidth level. New packet-switched MAN and WAN architectures and technologies are needed that are optimized for data traffic with carrier-class operational performance attributes.[28] Will this new architecture be Ethernet at the core metro or next-generation SONET with Ethernet handoffs? Planners need to understand the trade-off between efficiency, complexity, transparency, flexibility, manageability, and cost effectiveness. More specifically, metrics might include security, reliability, speed, throughput, access/backbone cost, equipment cost, sustainability, and margins (which also factor in operations/management costs). Figure 1-6 exemplifies such trade-offs.

Figure 1-6
An example of a multidimensional trade-off space for a carrier's deployment of broadband services.



Industry groups, such as the Institute of Electrical and Electronics Engineers (IEEE), are now developing new metropolitan fiber architectures with the stated goal of being able to handle data more efficiently than traditional SONET and to bring down the cost for service providers and companies. If successful, the groups will have a standard by 2003 that will define how the dual-ring architecture of SONET can be used without leaving one ring idle as backup the way SONET does. This new technology is known as Resilient Packet Ring (RPR). It takes advantage of the fact that data flows are less time dependent than voice. According to observers, these standards will let companies and service providers deploy metropolitan area optical bandwidth so that it can be distributed efficiently for IP, with a focus on Ethernet.

At the same time, there are major advances in SONET. These evolving systems are being called next-generation SONET. Vendors are now beginning to support the ITU X.86 and American National Standards Institute (ANSI) T1X1.5 Generic Framing Procedure (GFP) standard for direct Ethernet over SONET (EoS) services. This technology enables new products for LAN-to-LAN connections with fewer network elements.[29] Single-chip EoS mappers enable direct LAN-to-LAN communication over existing SONET/ SDH networks and are planned to be used in Ethernet switches, standalone data service unit/channel service unit (DSU/CSU) systems, and

SONET add/drop multiplexers (ADMs) within the LAN and metropolitan access (edge) network. EoS maps Ethernet frames into SONET/SDH payload and, as such, effectively transforms a portion of the SONET network into an invisible tunnel between LANs to provide TLS.

There is also interest in Generalized MPLS (GMPLS) because it extends the MPLS control plane to encompass time division (SONET ADMs), wavelength (optical lambdas), and spatial switching (incoming port or fiber to outgoing port or fiber). The extension of LAN capabilities to the metro area has also been called the optical Ethernet. Efforts of the IEEE 802.EFM, such as Ethernet in the First Mile (EFM), among others, address these developments. However, because of its field penetration, SONET/SDH will continue to play a key role for years to come, particularly if some of the newer technologies do not get deployed for any number of reasons. At the same time, vendors are adding features to the existing SONET platforms and, in doing so, they are bringing out next-generation SONET/SDH systems, which may play an important role in metro access/metro core deployments until (at least) 2005. The following section indicates some of the new technologies and specifications that are applicable to the metro space.

As noted, the idea of a carrier-supported LAN extension to a city level, that is, TLS service, is not such an esoteric concept as it is made out to be today. Bellcore's Switched Multimegabit Data Service (SMDS) and the IEEE 802.6 MAN standard were mid- to late-1980s carrier-grade broadband services. The former supported up to 34 Mbps and the latter supported 10 Mbps. TLS can be delivered via nonpacket architectures (such as SONET, next-generation SONET, and so on) as well as via packet architectures (such as GbE/VLAN, RPR, and so on). It is interesting to note that both SMDS and IEEE 802.6 services were packet based. Therefore, whatever intrinsic advantages packet brings to the table (as the strong RPR advocacy focuses on), these advantages are already a known quantity and already tried. If packet represents a breakthrough, the said breakthrough was already commercially available since the services of the 1980s. A discourse about the commercial success of packet-based TLS/MAN services such as VLAN, RPR, and so on should at least make reference to the commercial experience accrued at that time with SMDS and 802.6. Additionally, you should make note of the fact that efficient bandwidth packet technologies attempt to make the most parsimonious use of a supposedly plentiful commodity: bandwidth.[30] Why be so parsimonious (at the expense of added equipment costs/expenditures) with the resource whose price keeps dropping so rapidly? Furthermore, for at least a decade, providers of packet-based networks have regularly overengineered their networks by placing more bandwidth into play than necessary. This stratagem let them hide QoS issues from the user and keep customers as content as if they had utilized a private-line service.[31]

Some key new technology/specifications that are applicable to the metro access/metro core include the following:

- SONET/SDH and next-generation SONET/SDH
- GFP
- Optical Transport Network (OTN) and Intelligent Optical Network (ION)
- IEEE 802.3z GbE
- IEEE 802.3ae 10 GbE
- IEEE 802.1w Rapid Spanning Tree Protocol (RSTP)
- IEEE 802.17 RPR
- IEEE 802.EFM
- Optical Internetworking Forum (OIF) Optical User-Network Interface (O-UNI)[32]
- IETF GMPLS

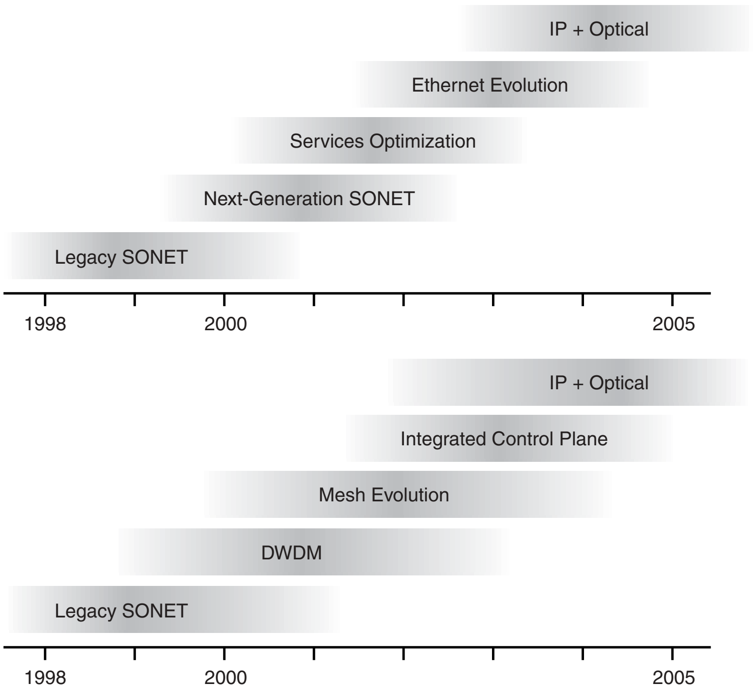
MAN evolution may include one or more of the following architectures and technologies:

- SONET/SDH
- Next-generation SONET/SDH
- IP over ATM over SONET
- IP over PPP over SONET
- DWDM (metro core)
- Wide WDM (WWDM), also known as coarse WDM (CWDM), for metro access
- Ethernet (Ethernet PHY in WAN) with VLAN and STP or RSTP
- Ethernet (10 GbE WAN PHY)
- RPR
- All-optical networking (future)

Figure 1-7 depicts one vendor's view of the possible evolution over the next few years.

As a consequence of this panoply of architectural choices, carriers will have to decide between next-generation SONET, metro DWDM, CWDM, optical Ethernet/GbE VLAN, optical Ethernet/10 GbE WAN, and optical

Figure 1-7
A Cisco view of the world.
Source: Cisco promotional materials



Metropolitan Network Migration
Traditionally, the metro has largely been supported by legacy SONET and SDH transport, with a TDM-based infrastructue. This structure limits scalability, is built with single-purpose equipment, and its centralized multistep provisioning is costly. Today, we're in the middle of the next-generation SONET revolution, bringing with it optically enabled IP, flexible architectures, and multiservice capabilities. This next-generation SONET environment is highly integrated for low cost and high density, guaranteeing simplified provisioning and integrated-packet services.

Ethernet/RPR. Each technology has advantages and disadvantages yet all will play a major role in the metro market's growth.[33] Some of the questions being posed include the following: Will metro DWDM solve the metro bandwidth bottleneck? Will next-generation SONET be able to handle the data-centric Internet world? Will easy-to-use, low-cost Ethernet work its way out of the access and into the metro core market? The transition of the lower layers, particularly the physical layer of the network, is a complex task that must be undertaken diligently because all higher layers supporting telecom and data communications are dependent on it. Figures 1-8 through 1-12 show the current SONET-based architectures, evolving incumbents' architectures, GbE/VLAN-based ELECs architectures, and two different evolving architectures for next-generation carriers.

Figure 1-8 Architecture evolutions: existing networks (subset).

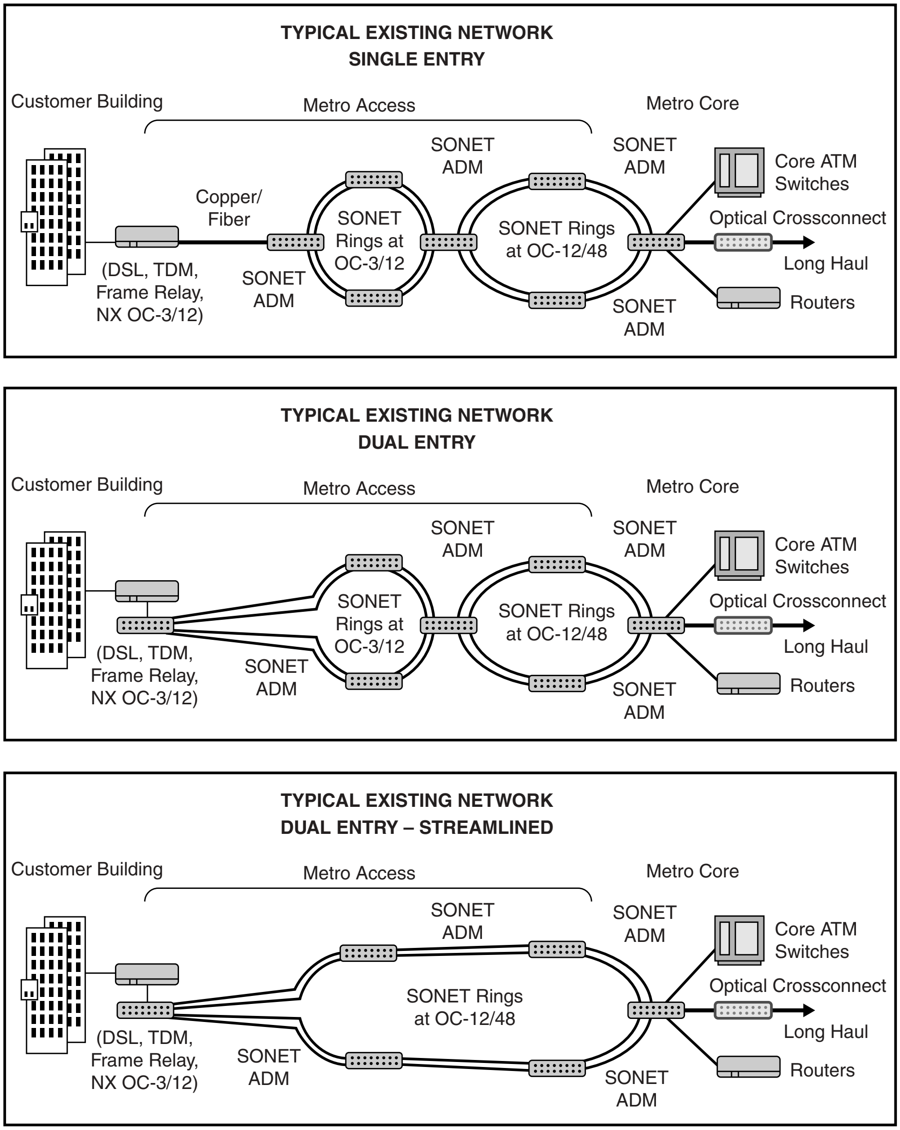


Figure 1-13 depicts at a macroscopic level the migrations that are occurring from the point of view of a protocol architecture (a more detailed protocol stack is included at the end of Chapter 3, “Traditional SONET Architectures”). Some of these migration steps are more likely than others:

- **IP over ATM to IP over SONET** ATM traditionally provided connection-oriented virtual channels. A mechanism exists in ATM for

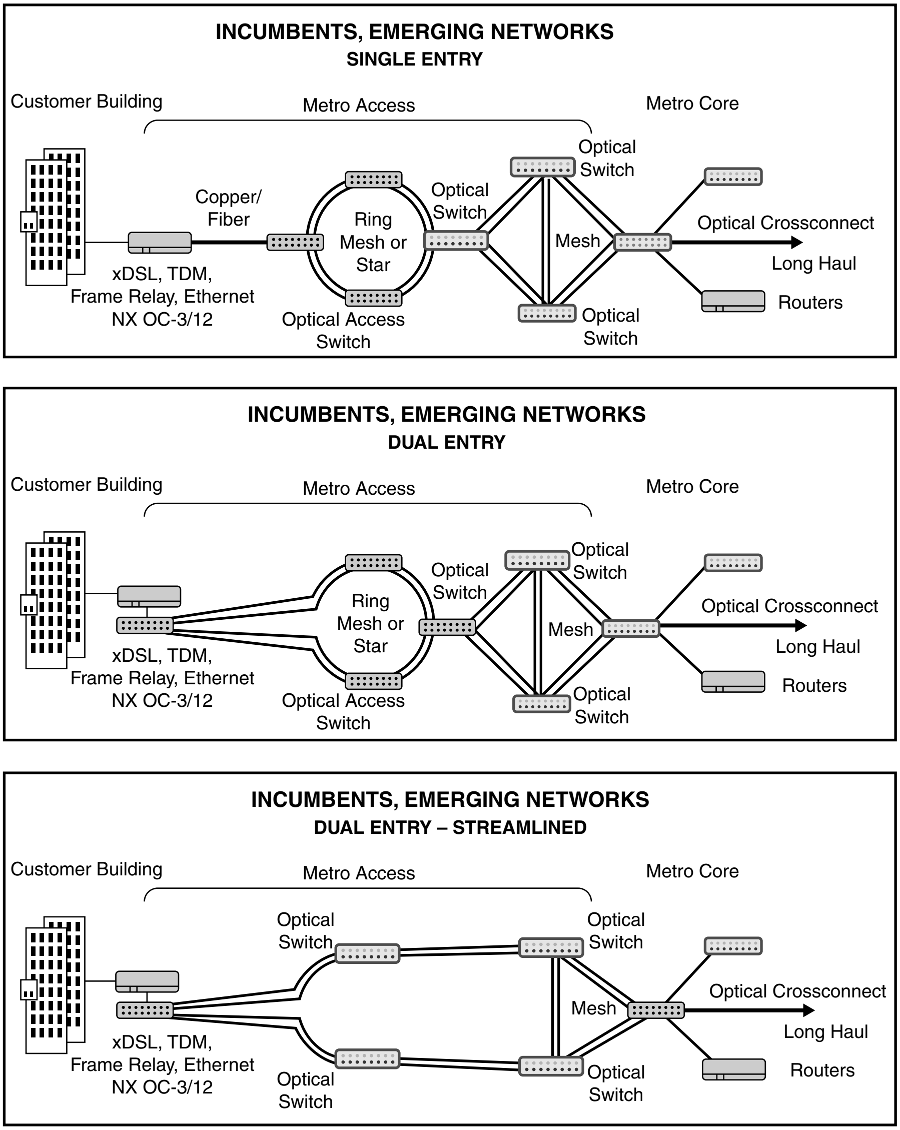
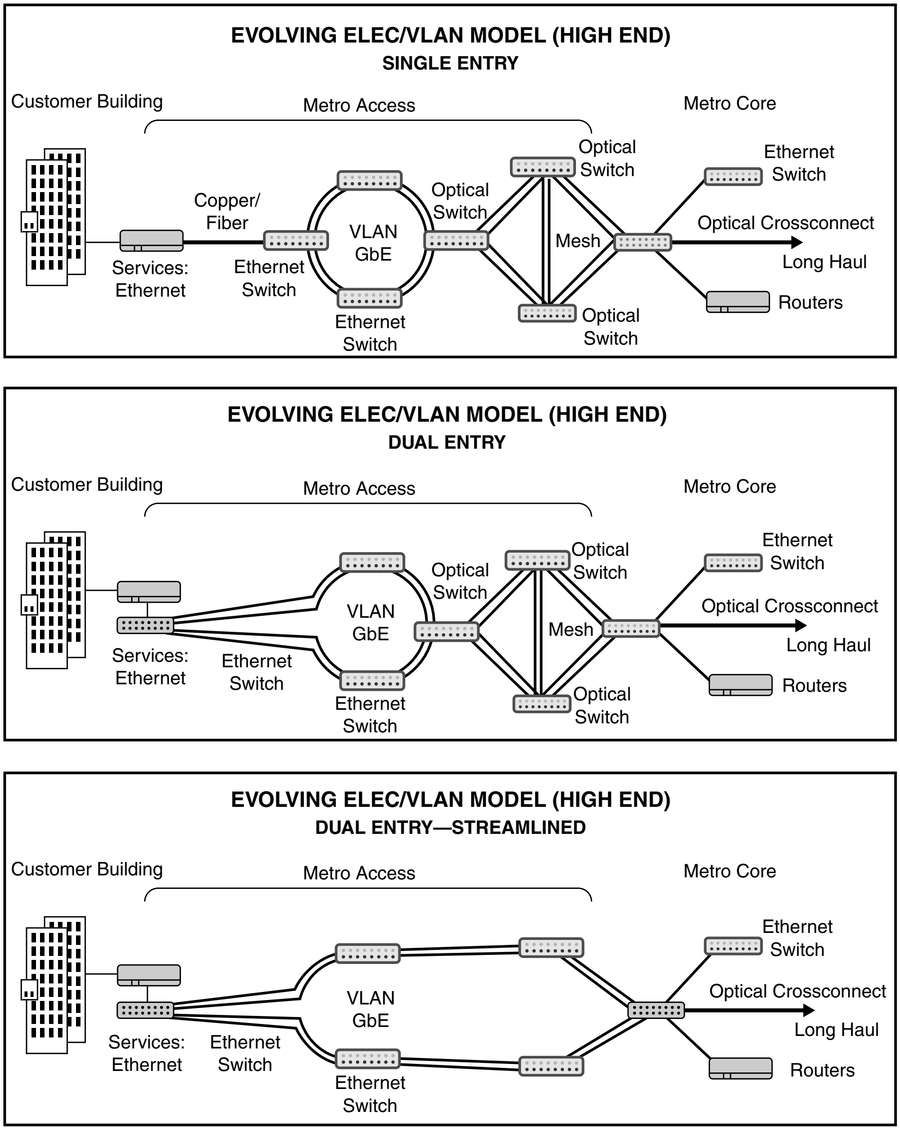


Figure 1-9 *Architecture evolutions: incumbents (subset).*

supporting QoS. SONET/SDH provides for TDM circuit transport and grooming. Protection and restoration (within 50 ms) are also provided by SONET. The OTN supports the transmission of one or more (SONET) optical channels over a fiber pair. When more than one channel is supported, a WDM technique is needed. For sometime now, ATM has been in the process of being replaced/eliminated from very

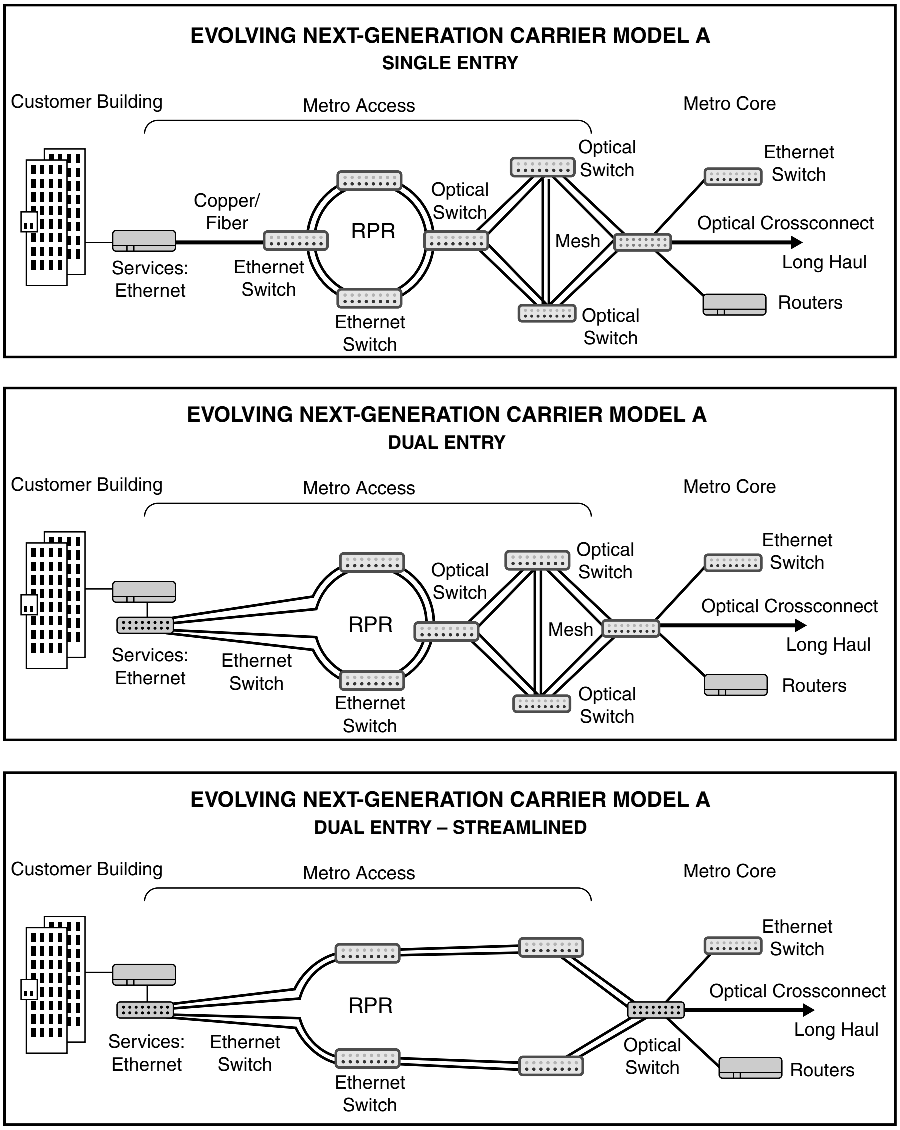
Figure 1-10 Architecture evolutions: GbE VLAN ELECs (subset).



high-speed connections (at 622 Mbps or higher). In this scenario, IP runs directly over SONET, which in turn runs on the OTN. MPLS has been added to the architecture to provide for connection-oriented services.

- **IP over SONET to IP over ION** In the next stage of the evolution, SONET could theoretically also be removed (although the SONET

Figure 1-11 *Architecture evolutions: next-generation carriers in packet mode (subset).*



framing is generally thought of as being retained), and IP could run directly over the OTN. In this scenario, the IP control plane could interact directly with the control plane of the OTN (an OTN with a dynamic control plane is often referred to as the ION) for dynamic establishment of optical connections. The eventual commercial outcome of this stage remains to be established, particularly with

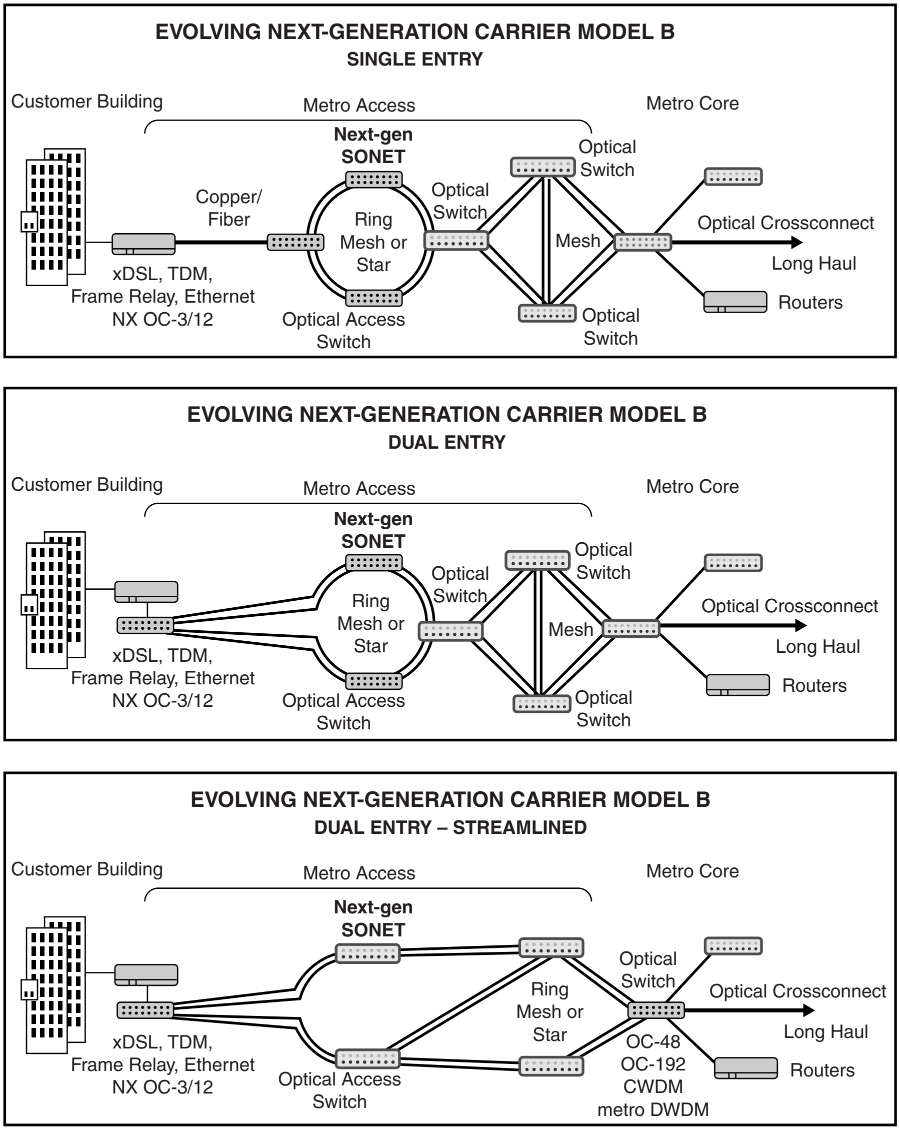


Figure 1-12 Architecture evolutions: next-generation carriers in circuit mode (subset).

regards to networks provided by carriers (the planner could always develop his or her own private network with full IP-to-ION interaction, but such a private network would likely be very expensive). The following are the factors affecting the commercial outcome: (a) full OAM&P capabilities are required by carriers; these capabilities are offered by SONET/SDH, but it is not clear that they

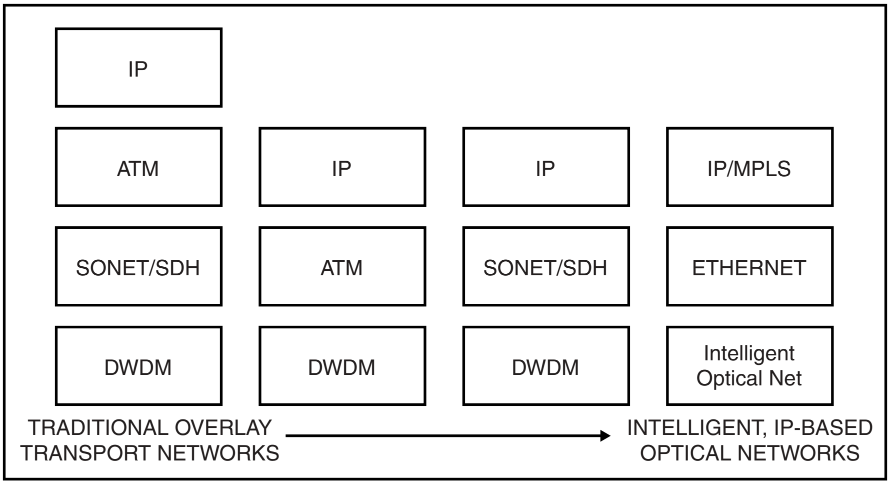


Figure 1-13
Macrolevel view of protocol architecture migration.

are offered by the ION; (b) the carriers may not want to relinquish provisioning control to users and switched virtual connections if either Frame Relay or ATM have *not* taken off; (c) the national network is not a simple cloud of a few nodes under the administrative domain of a single carrier as often implied by vendor presentations; and (d) only around 20,000 of the 4,600,000 commercial buildings in the United States currently have fiber, making the promise of instantaneous provisionability a mere chimera.

Table 1-2 maps some typical customer services to possible underlying technologies. It turns out that a multifaceted carrier will need more than a single technology and/or a single architecture. Table 1-3 lists the available metro access/metro core technologies and the advantages and disadvantages of each technology.

The following lists the metro optical network technology from an approach and supportive technologies standpoint. The key optical metro access/metro core approaches include

- Ethernet at the core
  - ELECs
- Next-generation SONET
  - Next-generation communications carriers
- DWDM/GMPLS
- All optical

Table 1-2

Best-Fit Mapping Between Services and Technologies

	CWDM/DWDM	TDM/SONET	VLAN/GbE	RPR
TLS	x	x	x	x
Private line	x	x		
Optical lambda	x			
Channel extension	x	x		
Multimedia		x	x	x

Table 1-3

Available Metro Access/Metro Core Technologies

Technology	Strengths	Weaknesses	Most Cost Effective
SONET	Standardized, proven protection, and performance monitoring	Inefficient for mixed traffic and service interface granularity	DS1, DS3, OC-3 LL; DS1, DS3 ATM; and DS1 Frame Relay
Hybrid ATM/SONET	Standardized, proven protection, and performance monitoring	Additional elements may be required, may increase network complexity, and have high costs	Frame Relay; DS3, OC-3 ATM; and transparent LAN
Multiservice DWDM	Highly scalable and supports multiple protocols	Still relatively expensive, often limited to optical multiplexing	OC-*x* LL; OC-*x* ATM; and GbE
Metro-optical Ethernet/IP	Low-cost efficient support of IP, native interface with Ethernet LANs and core routers	Does not support TDM without proprietary handling, single vendor solutions	Ethernet 10/100/1,000
Next-generation SONET	Very cost-effective utilization of metro bandwidth	Long-term scalability questionable, format and bit rate dependence	DS1, DS3, OC-*x* LL; DS1, DS3 ATM; DS1 Frame Relay; and STS-N Connect

Source: Geyser Networks
Note: Evolving technologies are not listed.
LL = long lines

The optical network technologies applicable at the metro level include

- DWDM
- CWDM
- DWDM filters
- Tunable lasers
- Optical switching
    - All-optical switching technology
    - Acousto-optics
    - Bubbles
    - Holograms
    - Liquid crystals
    - Liquid gratings
    - Micro-Electro Mechanical Systems (MEMS)
    - Thermo-optics
- Optical amplifier
- Optical crossconnects
- Advanced optical fibers and cables
- Arrayed waveguides and integrated optics
- Tunable filters

The following lists the kinds of nodal reliability for carrier-class network elements:

- **1:1 protection** A card protection scheme that pairs a working card with a protect card of the same type in an adjacent slot. If the working card fails, the traffic from the working card switches to the protect card. When the failure on the working card is resolved, traffic reverts back to the working card if this option is set. This protection scheme is specific to electrical cards.
- **1:*N* protection** An electrical-card protection scheme that enables a single card to protect several working cards. When the failure on the working card is resolved, traffic reverts back to the working card.
- **1+1 protection** An optical-card protection scheme that pairs a single working card with a single dedicated protect card.

The following lists developing standards and/or specifications of some key industry bodies that are pertinent to metro access/metro core networks:

- International Telecommunications Union (ITU)
- American National Standards Institute/Alliance for Telecommuniations Industry Solutions (ANSI/ATIS)
- Telecommunications Industry Association Fiber Optic (TIA FO)
- Full Service Access Network (FSAN)
- Bellcore/Telcordia
- European Telecommunications Standards Institute (ETSI)
- Optical Internetworking Forum (OIF)
- Institute of Electrical and Electronic Engineers (IEEE)
- Internet Engineering Task Force (IETF)

All of the activity by the leading bodies of the telecommunications and technology industry is certainly justified. Optical Ethernet and metropolitan networking companies have recently received billions in investment capital. During 2000 and 2001, venture capitalists funneled over $3 billion into 37 startups in the optical Ethernet space.[34] The biggest portion of this funding (about $1.67 billion) has gone to companies developing metropolitan area networking equipment for use by enterprise customers and carriers. Another $1.54 billion has gone to service providers specializing in optical Ethernet connectivity. A smaller portion (about $415 million) has gone to firms supplying components to equipment vendors. Much of the investment in optical Ethernet has come from leading firms, such as Goldman Sachs & Co., Intel Capital, New Enterprise Associates, and WorldView Technology Partners. Each of these firms has backed at least five separate companies. Intel has been the most active in this realm, having invested in seven separate optical Ethernet or metro optical transport companies.

1.3 Course of Investigation

In this book we examine traditional-but-fast-improving as well as evolving architectures. Architectures, technologies, and services for high-end commercial applications are explored in detail. A companion book by these authors, *Ethernet-Based Metro Area Networks*, lays the groundwork for considering SONET—in its traditional and next-generation forms—in contrast to and comparison with Ethernet-based wide area architectures such as VLANs, Rapid Spanning Tree Algorithm, RPRs, 10 GbE, and so on. Our purpose is in part to evaluate the latter against multilambda ring solutions. Ring versus mesh architectures are also studied.

In this volume we start with the assertion that significant advances in optics and next-generation SONET enable features and services that can rival what the new Ethernet architectures offer. This hypothesis is tested by examining the performance of the technology in the context of operating systems. Both books pose the question of whether Ethernet-based architecture can scale in size, robustness, reliability, and OAM&P capabilities as well as enhanced SONET systems, but they do so from two different perspectives. In this volume, we cover the following topics.

This chapter sets the stage for an assessment of the advances in SONET systems that enable them to compete with new architectures. Chapter 2, “Optics Primer,” provides a short primer on optical technology. Chapter 3, “Traditional SONET Architectures,” looks at traditional broadband architectures, such as SONET/SDH, focusing on the self-healing robustness and OAM&P features. Chapter 4, “Next-Generation SONET and Optical Architectures,” looks at next-generation SONET and the advantages it affords; a number of next-generation carriers are going this route to provide advanced services. Chapter 5, “Dense Wavelength Division Multiplexing (DWDM) Systems,” focuses on optics and DWDM, a technology so far optimized for long-haul use. Chapter 6, “Coarse Wavelength Division Multiplexing (CWDM) Technology,” looks at a number of other practical applicable technologies for TLS delivery, such as CWDM. Chapter 7, “All-Optical Networks,” covers evolving all-optical/intelligent optical-networking technologies. Chapter 8, “GMPLS in Optical Networks,” discusses wavelength services managed by GMPLS. Chapter 9, “Free-Space Optics (FSO),” discusses FSO, which some people are proposing as a metro access technology. Finally, Chapter 10, “Practical Design and Economic Parameters,” looks at practical design approaches and economics to validate the premise that these evolving technologies and architectures (both the Ethernet and the next-generation SONET) can be reliably and sustainably deployed in the field.

How is the market climate as operators everywhere confront these or similar questions? The following are some indicators from the trade press and popular writings of late 2001:[10]

The optical revolution deferred It now appears that deployed optical networks will not be moving from 10-Gbps rates to 40-Gbps rates any time soon. This is a big shock in an IT world that is accustomed to the regular pace of Moore’s Law, which yielded faster processors, larger memories, and bigger disks. At the same time, the shift from optical-to-electrical-to-optical (O-E-O) conversion to an all-optical network has been delayed significantly. Also, the penetration of WDM technology from the WAN to the metro area has not yet happened as ILECs concluded they had enough fiber and most CLECs have left the scene. Will there be a lost generation of optical component vendors and system companies that were founded on the premise that we needed 40 Gbps, lots of wavelengths, and pure optical transport and switching?

Disappointments in broadband access The DSL industry has suffered a setback because most of the DLECs have gone out of business. Cable Multiple System Operators (MSOs) are moving ahead with broadband deployment, but that does not help most businesses. Innovations for optical access—including PONs, Ethernet over fiber, and lower-cost SONET—are all promising, but have not yet been widely deployed, in part because the incumbent carriers have cut back on spending and postponed new initiatives. Many industry participants were counting on a massive adoption of low-cost broadband access to enable new multimedia services and the transformation of the Web into a platform for more commerce and entertainment. Another sure thing was that everything was moving from the enterprise into the Net—web servers, storage, applications, the works. The fall of Exodus, the missteps of many ASPs, and the perilous financial state of some Internet data centers have caused many enterprises to rethink their strategies.

Deferral of packet voice and MPLS It has been an axiom in our business that everything is moving to IP, including voice. Although there has certainly been more progress in the past year, the widely anticipated shift to packet voice is not apparent in the networks of the major carriers. In fact, the whole premise of multiservices networking seems to be given lip service only. Where are the MPLS networks that we were expecting to be carrying multiservices traffic by now? Again, the vendors have developed the tools, but the service providers have not yet seized the opportunity to offer next-generation services.

Areas where innovation appears to be taking hold: **metro service providers** Although many segments in the service provider industry have been experiencing difficulties, new metropolitan carriers are blazing a new path with breakthrough technologies and innovative pricing. They are building remarkable optical IP Internet infrastructures within key metropolitan areas domestically and internationally. By offering virtually unlimited, unmetered metro area communications at a fixed price, these revolutionaries are eliminating the bandwidth barrier. One provider offers 100 Mbps of nonoversubscribed Internet access for a flat $1,000 per month. This makes access to the Internet and metropolitan service seem like simple extensions to a LAN. Others offer high-speed, low-cost interchange between last-mile providers and the Internet.

1.3.1 Our Assumptions

The assumptions behind this investigation can be summarized as follows: New architectural models are emerging and need to be considered by carriers. There may be advantages in these new models and technologies. These architectures include GbE/10-GbE-based backbones (optical Ethernet) using VLAN techniques, SONET and next-generation SONET (where SONET supports native Ethernet interfaces to the user to achieve cost savings over traditional SONET), metro-optimized WDM (CWDM) possibly under GMPLS control, and evolving RPR. Customers may prefer an Ethernet handoff for carrier services because they are interested in TLS. They also want carrier-class, highly reliable, and secure services, especially if they have to funnel most of their corporate data traffic plus voice over a small set of multimegabit and even possibly multigigabit circuits.

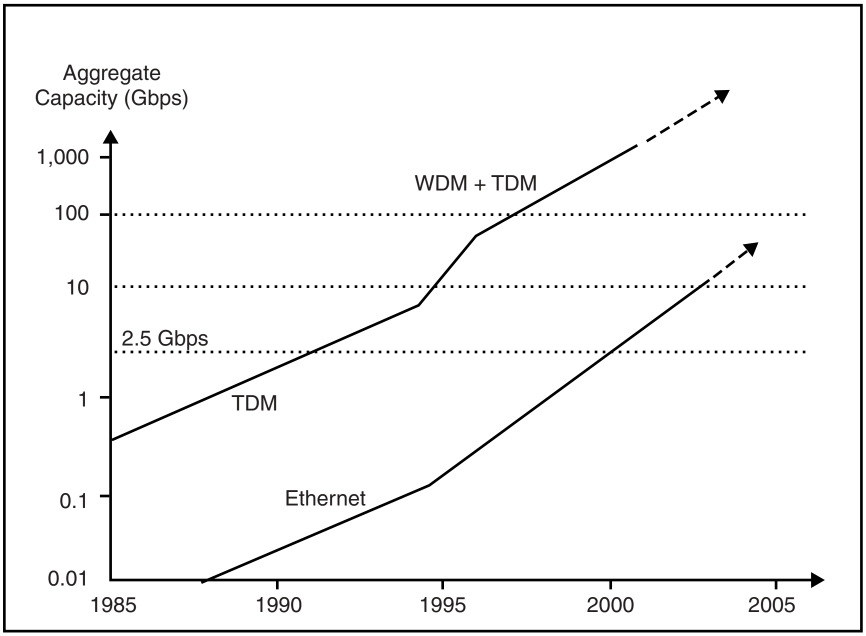
There is already excessive hype on the new GbE/10-GbE-based technologies for metro applications. They are being offered as a panacea by proponents. The economics presented are very superficial and miss the true test of what it takes to run a carrier function. There is an excessive selling of technology rather than services. Not enough true consideration has been given to robustness, scalability, and provisionability. Some of the new players are naive and do not have a sustainable business model for offering a single-platform, single-service proposition, leading to low market penetration, revenue, net income, valuations, and return for investors but, these are making up said lacunae by the might of high-gear public relations. Some recent Chapter 11 events substantiate this assertion.

We believe that the next-generation players who will be successful are those who have a panoply of reliable and secure carrier-class services. These carriers have to be multiarchitecture, multitechnology, multiservice, and multi-QoS oriented. These carriers sell services, not technology. These carriers give due consideration to developing technologies whether they are brand new or timely enhancements of old standbys, and use all of them to pursue a sophisticated technical and financial analysis of their advantages.

1.4 A Short Synopsis of Baseline Technologies

SONET and WDM are two key technologies of interest. Some basic preliminary information on these two technologies is included. Figure 1-14 depicts

Figure 1-14
Capacity of TDM, TDM+WDM, and Ethernet systems over the years.



the increase in aggregate capacity over the years of TDM (SONET/SDH) and TDM+WDM systems as well as the increase in Ethernet speeds.

1.4.1 SONET and Next-Generation SONET Positioning

SONET/SDH is a standard (the former from ANSI and the latter from ITU) intended to support the interworking of optical equipment at various speeds, typically 2.5 Gbps, 10 Gbps, and in the near future 40 Gbps (lower speeds are also supportable). Interworking is achieved by defining a set of optical parameters, data-framing formats (the frame format for the basic building block in shown in Figure 1-15), clocking speeds, and OAM&P mechanisms. Figure 1-16 depicts the basic SONET arrangement, whereas Figure 1-17 illustrates the point-to-point environment of a SONET link.

In the past few years, manufacturers and vendors of IP-based equipment have claimed that the inefficiencies of TDM-based technology could not accommodate the growth of Internet and data traffic. They suggested that

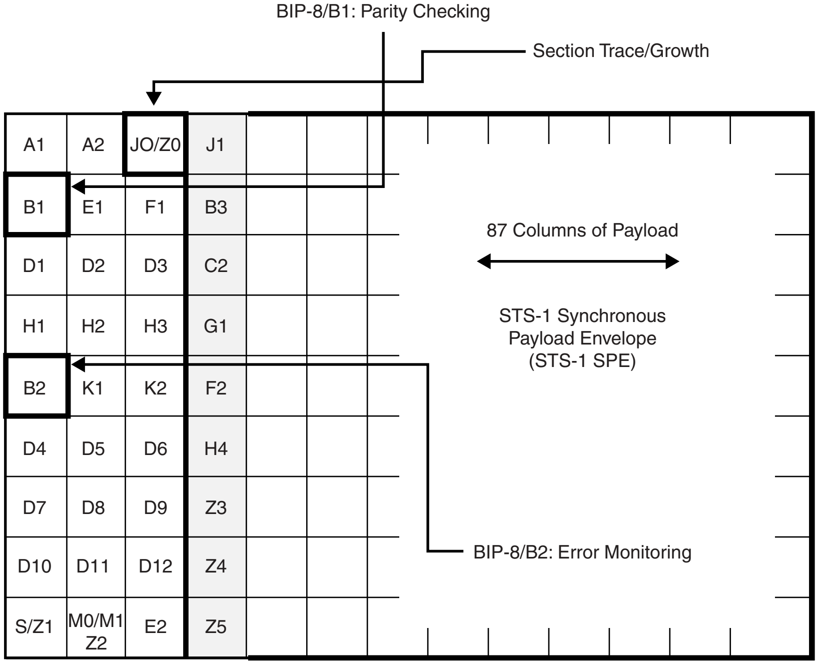


Figure 1-15
SONET's path, line, and section.

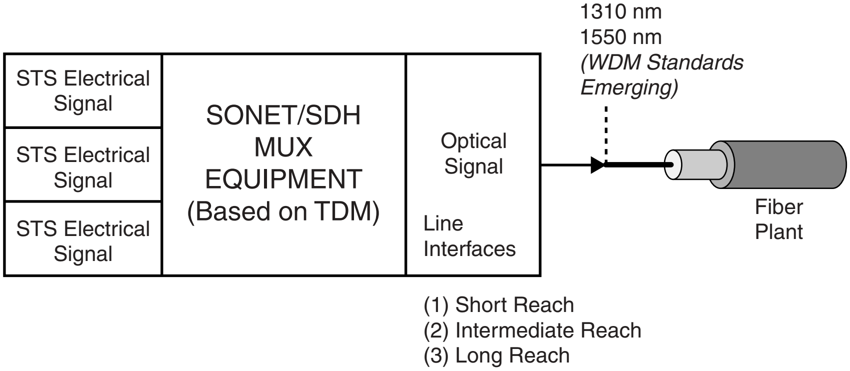
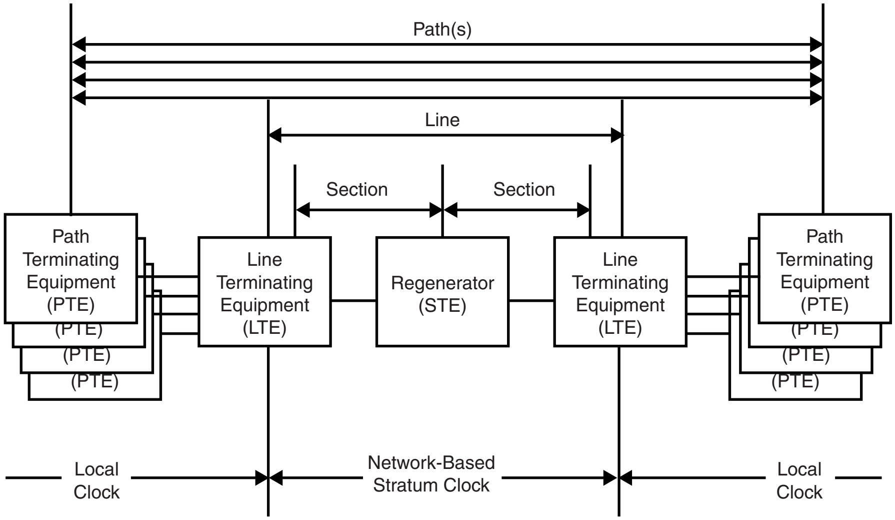


Figure 1-16
SONET environment and drivers.

- Standards based: Allows multivendor interworking
- Leveraged emerging fiber-optic transport technologies
- Synchronous transmission: Addresses timing variations
- Reduces complexity of operations: Automation of OAM&P
- Network scalability: Enables multiplexing of lower-speed streams onto single high-speed transport

Figure 1-17
SONET Path, Line, and Section

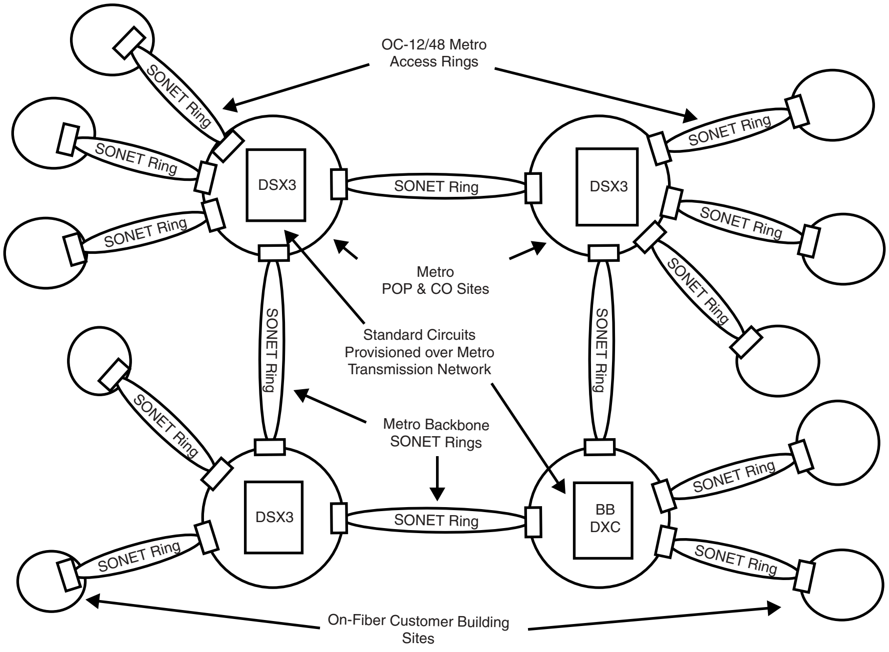


if correct, this analysis would drive the SONET metro market to flatten and, in a few short years, disappear completely. But that hasn't happened. Analysts are forecasting that the SONET metro market will grow from around $5 billion in 2001 to more than $9 billion by 2003. However, there is one major caveat: According to proponents, the SONET equipment of the early to mid-1990s is nothing like the next generation of sophisticated, multiservice, metropolitan optical nodes that will dominate this space in the coming years.[33]

SONET/SDH metro access architectures favor a collection of two node rings instead of multinode rings, as illustrated in Figure 1-18.[35] SONET/SDH rings are utilized because of the high-speed restoration in case of a fiber cut or other link anomalies. Two node rings are often used at the metro core level between POPs because of the capability to support maintenance on transmission systems without affecting more than one site. (Multinode rings are typically not used in local environments because any upgrade that is done to one access site requires an upgrade to all of the systems on the ring, instead of only the site that needs it.) SONET/SDH multinode rings are more often used in long-haul backbone transport infrastructure. Multinode rings cost more than mesh networks, but they provide faster restoration where traffic restoration over distances is more of an issue. For a typical regional architecture, see Figure 1-19.

This topic will be treated at length in Chapter 3.

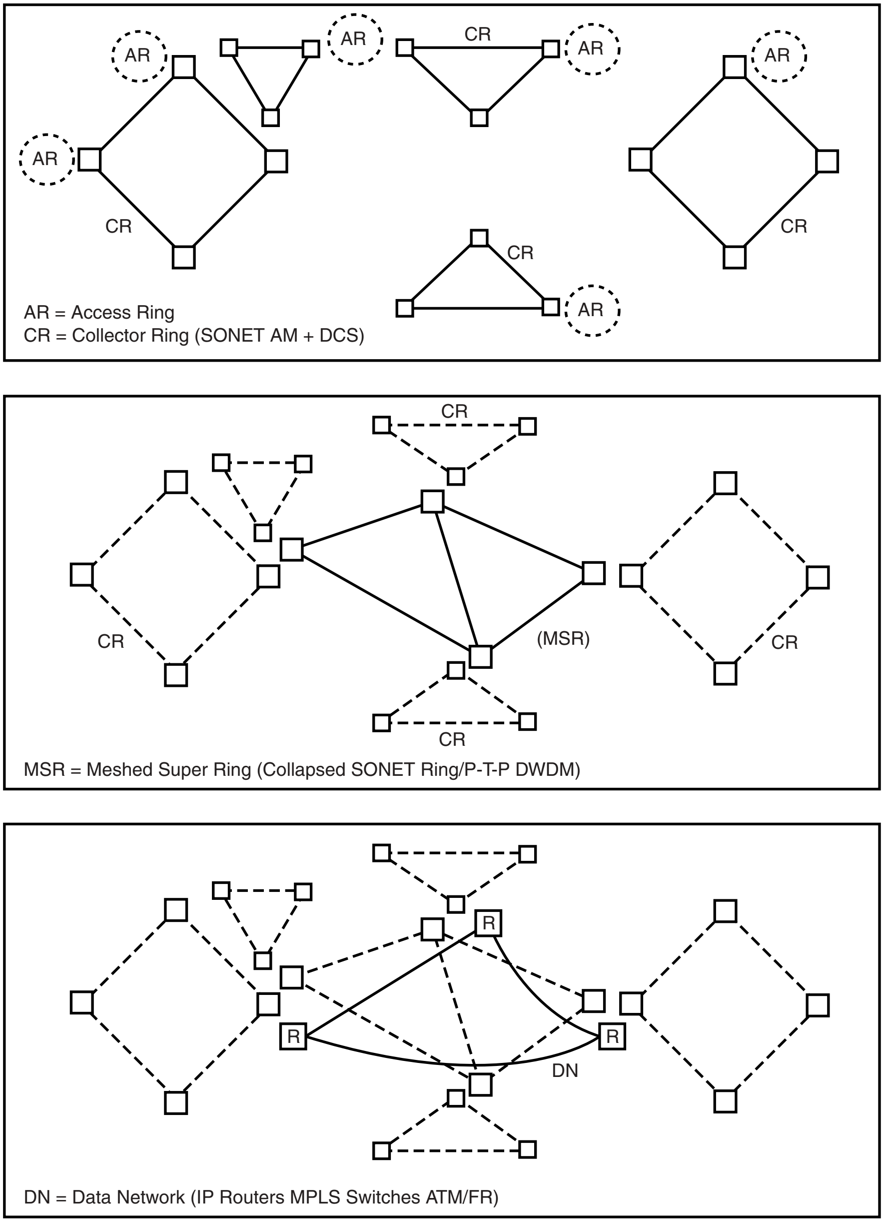
Figure 1-18 Typical SONET/SDH local architecture.



1.4.2 DWDM Positioning

DWDM systems are used to carry multiple signals on a single fiber simultaneously. More than 16 million lambda miles of DWDM systems have been installed across the globe (a Nortel Networks statistic). Modulation on a fiber in a traditional system is unipolar amplitude shift keying (ASK), which is another name for amplitude modulation (AM). In the optics world, intensity modulation (IM) describes ASK/AM. Traditional systems ASK modulate a light beam at, for example, 1,310 nm. In a simple modulation environment like IM, the clock speed determines the digital throughput of the system; for example, if the clock is 2.4 GHz, then the bit rate is 2.4 Gbps. Because the optical domain is large (the infrared spectrum where telecommunications optics operate has a bandwidth of 10^{14} Hz), you should theoretically be able to derive between 10^{14} bps and 10^{10} bps; this is between 0.5×10^5 (50,000) and 0.5×10^6 (500,000) times more achievable on a system with a single ASK-modulated signal operating at some fixed frequency.[36] Hence, if you were to employ IM-based systems operating at practically spaced spectral points, you should be able to obtain additional throughput. In effect, you want to use frequency division multiplexing (FDM)

Figure 1-19 Typical regional architecture.



techniques in conjunction with the digital ASK-based modulation employed at each of the frequencies. This methodology is WDM (or DWDM when 16 or more frequencies are used). WDM technology started with systems that carried 2, then 4, then 8, then 16, then 32, and so on light beams. Systems with more than 160 channels are now available. For example, some systems could support 64 channels unprotected or 32 channels protected. A 16-channel system in essence provides a virtual 16-fiber cable with each frequency channel serving as a unique STM-16/OC-48 carrier. Advances in aggregate bandwidth have been seen in recent years, from 400 Gbps on high-end systems evolving to 800 Gbps. Near-term projections are in the terabit per second range. You can also observe DWDM channel growth. Although the current upper range is OC-192 (10 Gbps), the near-term projected upper range is OC-768 (40 Gbps). The availability of mature supporting technologies, such as precise demultiplexers and erbium-doped fiber amplifiers (EDFAs), has enabled DWDM with 8, 16, and even higher channel counts to be commercially delivered. DWDM systems have been used extensively in the long-haul space and typically cost from $100,000 to $1 million or more.

There are active and passive elements in a DWDM system. Figure 1-20 illustrates the active and passive components that comprise a DWDM system. Active components create the optical signals (wavelengths of light) that travel throughout the fiber-optic network. Passive components manipulate and guide light along the fiber.

Transmitter modules (TxM) are responsible for sending input signals, whereas receiver modules (RxM) are responsible for receiving output signals. Each wavelength enters and leaves the DWDM system separately through TxM and RxM. As indicated in Figure 1-21, TxM and RxM consist of passive and active components. The passive components (isolators, couplers, attenuators, detectors, and DWDM mux and demux components) are used to route and guide the wavelengths through the network. The active components (modulators, transmission lasers, and wavelength lockers) are used to create, modulate, or amplify the wavelengths. TxMs and RxMs can

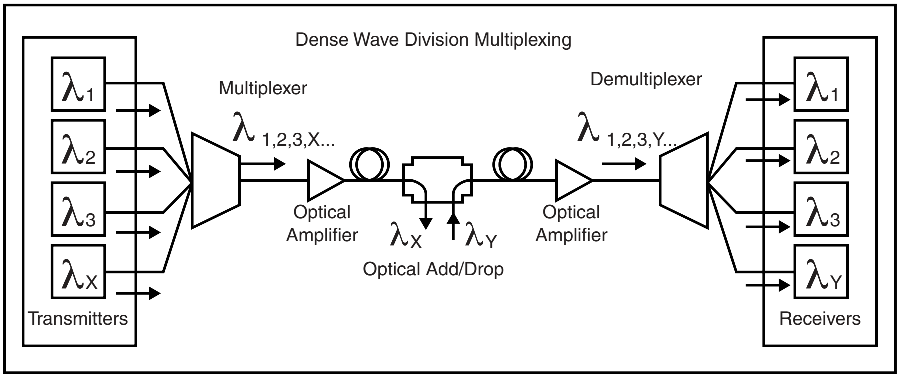


Figure 1-20
Active and passive components of a DWDM system.

Figure 1-21
Transmitter and receiver components of a DWDM system.

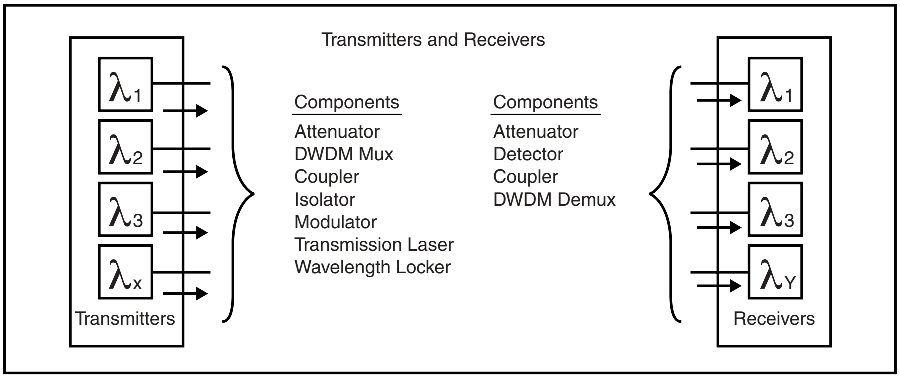
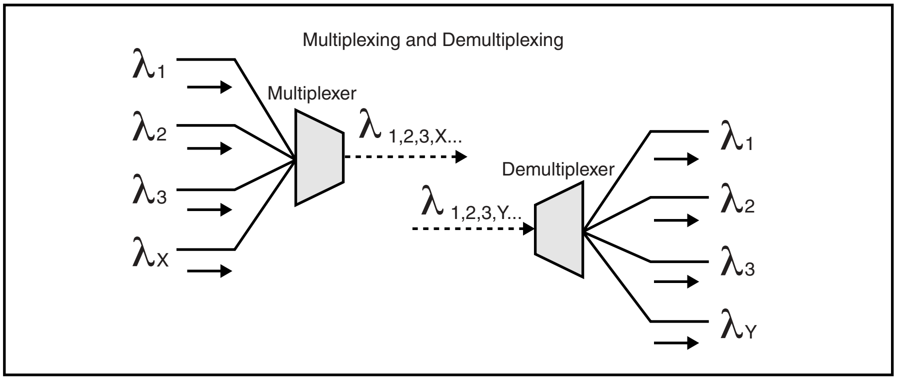


Figure 1-22
DWDM multiplexing/ demultiplexing.



also be packaged together as a transceiver. A transceiver transports input and output light signals through the system, whereas TxM and RxM can only transport light in one direction. The most common form of DWDM uses a fiber pair—one for transmission and one for reception. Systems exist in which a single fiber is used for bidirectional traffic, but these configurations must sacrifice some fiber capacity by setting aside a guard band to prevent channel mixing. They also degrade amplifier performance and reliability.

DWDM multiplexing and demultiplexing (see Figure 1-22) requires separating and combing wavelengths. There are three primary technologies used to perform the multiplexing and demultiplexing functions: thin-film filters, fiber Bragg gratings (FBGs), and arrayed waveguides.

Although WDM has been a widely known technology for several years (a little more than FDM by another name), its early application was restricted to providing two or four widely separated wideband wavelengths. During the mid- to late-1990s, the technology evolved to the point where parallel wavelengths could be densely packed and integrated into a transmission system with multiple signals in the 192- to 200-THz range. There are economic advantages to using DWDM when the transmission costs are high, such as in long-haul applications. Such use is justified by standard transmission-versus-multiplexing cost calculations. For example, transmitting 40 Gbps over 600 km using a traditional system would require 16 separate fiber pairs with regenerators placed every 35 km for a total of 272 regenerators. A 16-channel DWDM system, on the other hand, uses a single-fiber pair and 4 amplifiers positioned every 120 km for a total of 600 km.

This topic will be treated at length in Chapter 5.

1.5 Key Metro Access and Metro Core Players

The following lists show some of the key players in the services and technology broadband arenas at the time of this writing. Although lists such as these are subject to change over time, they do provide a snapshot of the industry.

Examples of carriers with broadband opportunities include the following:

- **ILEC metro optical networks** Bell Canada, BellSouth, Qwest, SBC Communications, and Verizon
- **Next-generation communication carrier metro optical networks** To be determined
- **ELEC metro networks** Cogent, Telseon, and Yipes
- **CLEC metro networks** Digital Teleport, Inc., Electric Lightwave, Time Warner Telecom, and WorldCom
- **Interexchange Carrier (IXC) and utility metro optical networks** AT&T, AT&T Canada, Cable & Wireless USA, Enron Communications, Global Crossing, Level3 Communications, MetroMedia Fiber Networks, Qwest, Sprint Corporation, and Williams Communications
- **European carriers** British Telecom, Cable & Wireless, Deutsche Telekom, France Telecom, Hermes Europe Railtel/GTS Trans-European Network, RegioNet, Tele Danmark, Telecom Italia, Telemonde, Telenor, Viatel, and WorldCom

- **Storage services providers (SSPs)** Arsenal Digital Solutions Worldwide, Inc., Centripetal, Inc., ManagedStorage International, Sanrise, Storability, Inc., StorageNetworks, StorageWay, and WorldStor

The following are examples of vendors of optical/broadband equipment:

- Component vendors
  - Agility
  - Alcatel Optronics SA
  - AMP
  - Amphenol Fiber Optic
  - Anritsu Corp.
  - Aurora Photonics, Inc.
  - Blue Sky Research, Inc.
  - Bragg Photonics, Inc.
  - Burleigh Instruments, Inc.
  - Calient Networks
  - Corning
  - DiCon Fiberoptics
  - GN Nettest
  - Integrated Photonic Technology (IPITEK)
  - JDS Uniphase Corp.
  - Mitel Semiconductor
  - Optical Micro-Machines, Inc. (OMM)
  - Optilas International
  - Precision Optics Corp. (POCI)
  - Southampton Photonics (SPI)
  - Tektronix
  - Wavesplitter Technologies
- Equipment vendors
  - Agere Systems
  - Alcatel Network System
  - Alidian Networks
  - APA Optics

  - Avanex
  - Bookham Technology
  - Canoga Perkins
  - Ditech Communications
  - Dynarc, Inc.
  - Ericsson
  - Hewlett-Packard
  - Hitachi Cable America
  - Lightchip
  - Lightwave Microsystems
  - Nortel Optoelectronics
  - Optical Switch Corp.
  - Radiant Photonics

- Metro DWDM/next-generation SONET vendors
  - ADVA
  - Alcatel USA
  - Ciena
  - Cisco
  - Ericsson
  - Fujitsu Network Communications, Inc.
  - Kestrel Solutions
  - Lucent Technologies
  - LuxN, Inc.
  - Marconi Communications
  - Metro-Optix
  - NEC Networks
  - Nortel Networks
  - Optical Networks, Inc. (ONI Systems)
  - Redback Networks
  - Siemens
  - Sorrento Networks
  - Sycamore Network, Inc.

  - Tellabs
  - White Rock Networks
  - Zaffire, Inc.

- Vendors of access-based metro solutions
  - Alloptic
  - Appian Communications
  - Atrica Inc.
  - Auro Networks Inc.
  - Extreme Networks
  - Force 10 Networks, Inc.
  - Jedai Broadband Networks
  - Lantern Communications Inc.
  - Luminous Networks Inc.
  - Native Networks
  - Nortel Networks
  - Riverstone Networks
  - Tropic Networks Inc.
  - VIPswitch
  - World Wide Packets

- Vendors of optical packet nodes
  - AcceLight Networks
  - Atoga Systems
  - Lightscape Networks
  - Village Networks

- Vendors of storage area networks
  - Adva
  - Akara
  - Brocade Communications Systems
  - CNT Networks
  - EMC Corp.
  - Entrada Networks
  - Finisar Corp.

- Gadzoox Networks, Inc.
- INRANGE Technologies Corp.
- Lucent Technologies
- LuxN Inc.
- McDATA Corp.
- Network Appliance Inc.
- Nishan Systems
- Nortel Networks
- ONI Systems
- QLogic Corp.
- SAN Valley Systems
- Storage Computer Corp.
- Vixel Corp.
- YottaYotta Networks

1.6 Rosy or Red All Over?

Even as this book examines next-generation SONET systems, vendors and proponents are now inundating the trade press, conferences, customers, and venture capitalists with an avalanche of claims about the payback of Ethernet-based metro systems. Table 1-4 is typical, and it appears in a published advocacy white paper from a very reputable source. The table shows, from a bottom-line standpoint, a purported rosy picture of payback when using Ethernet-based metro systems (for example, the internal rate of return [IRR] is 87.6 percent, the earnings before interest, taxes, depreciation, and amortization (EBITDA) are positive in year 1, and so on). Is this table really rosy or is it red all over?

The reality is completely different. The watershed years of 2001 and 2002 with dozens of failures in the telecom sector, including CLECs (GST and ICG), RLECs (Telergy and Winstar), fiber IXCs (360Networks), ELECs (Yipes), and ISPs (PSI Net and Exodus), and the dot-com space corroborate the position that it takes quite a long time to turn EBITDA and net income positive. A highly skilled telecom team might achieve that in three years. The CLECs were fond of perturbing the textbook definition of gross margin (revenues minus all expenses with the single exception of sales, general, and administrative [SGA]), and always claimed a gross margin of around

Table 1-4

Advocacy White Paper from a Very Reputable Source

	Pro Forma Income Statement (Deployment in 10 Cities)				
	2001	**2002**	**2003**	**2004**	**2005**
Revenues	$129,700,000	$265,100,000	$409,500,000	$594,800,000	$717,400,000
Expense items					
OAM&P	($38,900,000)	($79,500,000)	($122,800,000)	($178,400,000)	($215,200,000)
Sales and marketing	($36,300,000)	($55,700,000)	($71,700,000)	($104,100,000)	($125,500,000)
General and administrative	($15,600,000)	($23,900,000)	($30,700,000)	($44,600,000)	($53,800,000)
Leased fiber	($17,300,000)	($34,800,000)	($50,800,000)	($72,400,000)	($81,900,000)
Building lease	($5,500,000)	($9,700,000)	($12,100,000)	($16,000,000)	($16,000,000)
Total expenses	($113,600,000)	($203,500,000)	($288,100,000)	($415,600,000)	($492,500,000)
EBITDA	$16,200,000	$61,600,000	$121,400,000	$179,200,000	$225,000,000
Depreciation/amortization	$11,400,000	($15,600,000)	($21,100,000)	($30,700,000)	($39,100,000)
EBIT	$4,800,000	$46,000,000	$100,200,000	$148,500,000	$185,800,000
Interest	($1,700,000)	($1,900,000)	($2,200,000)	($2,900,000)	($3,200,000)
Taxes	($1,900,000)	($18,400,000)	($40,100,000)	($59,400,000)	($74,300,000)
Net income	$1,200,000	$25,700,000	$57,900,000	$86,200,000	$108,300,000
Operating cash flow	$12,600,000	$41,300,000	$79,100,000	$116,900,000	$147,500,000
Capital	($57,100,000)	($20,700,000)	($27,900,000)	($47,900,000)	($42,000,000)
Net cash flow	($44,500,000)	$20,500,000	$51,200,000	$69,000,000	$105,400,000
Discount rate	12%				
Net present value (NPV)	$116,800,000				
IRR	87.6%				

80 percent. A highly skilled telecom team might achieve a 40 percent gross margin. A net telecom income of 20 percent, accumulated over a five-year startup horizon, is a major accomplishment. At the end of 2001, the leading ELEC had received over $300 million in funding and had achieved, after approximately three years of existence, an annualized income of $25 million, or less than $0.10 for every dollar invested.

Telecommunications carriers are utilities, and net incomes in the 20 percent range are major accomplishments.

A number of carriers resorted to questionable bandwidth swaps and/or wholesale business to show any revenue numbers.

Although Table 1-4 reflects Ethernet-based technology and not SONET or WDM technology, there is no silver bullet on the horizon. As we discussed in our companion book (from which the rest of this section is based), the major cost components are in the building fiber access spur, the operations cost, and the SGA, not the network element in the network, much less the protocol engine.

For protocols, architectures, or Request for Comments (RFCs) to see deployment, they must address and solve the real underlying problem at hand. Perhaps what the industry needs to do is not just build some new architecture, but also tackle the issue related to the financial viability of delivering tangible services to users within an acceptable provisioning interval. Real-life analysis demonstrates that the most expensive element of delivering said broadband service (unblocked/unoversubscribed 10-, 100-, 1,000- or 10,000-Mbps throughput) is not in the equipment itself when it is selected appropriately or the backbone ring when it is designed correctly (except for the very high-end Gbps/10-Gbps cases), but in the fiber access to the building from the backbone itself running perhaps just in the street below or maybe a block or two away. This portion of the network is called a *spur*. The building's riser(s) also can be significantly expensive (whether as a one-time cost or as a monthly fee back to the building owners).

The spur cost is typically in the $75,000 to $200,000 range (at times even more). When this cost is amortized over 30, 40, 50, or 60 months, the monthly recurring charge (MRC) just equating to the spur is in the $1,300 ($75,000/60) to $7,000 ($200,000/30) range, with $1,500 being typical. This figure is high, particularly because a carrier may be able to find only one to three broadband-paying customers in a building. We do not believe that anyone can find 67 such customers in a building, as some first-generation (1G) ELECs have stated that they need. ELECs are typically characterized by being single-architecture, single-service, and single-technology CLECs that rely, in most instances, exclusively on shared GbE metro backbones. They follow the trajectory of those companies in the early 1990s that deployed shared (Fiber-Distributed Data Interface [FDDI]) backbones to provide services. If none come to mind, it is because they are all long gone.

A small subset of analysts do see what the issues are, as in the following quote: "Optical access will be on top 10 lists for the next few years, if only because it takes so long to build out 'last mile' networks, especially with fiber. Fiber has been deployed in enormous quantities in the core of national networks, and in the past few years companies and a host of utilities and municipalities have made significant strides in getting fiber down the busy streets of financial districts. Getting fiber to that last mile is a pain, and it is expensive. Yet the emerging market for broadband services from small- and medium-sized businesses insists that fiber be drawn as nearly as possible to the end user, either to the building basement or the wiring closet. Getting fiber that close requires operators to spend lots of money on installation, so end equipment must be low-cost, scaleable, and remotely manageable."[37]

Figure 1-23, which is compiled from real carrier studies, points to the underlying problem that has to be solved before the industry can effectively move forward and deliver services in a sustainable, EBITDA-positive, and preferably net-profit-line-positive manner. Right now, when the largest group of customers want in the range of 10- to 100-Mbps throughput per building (considering the million or so Frame Relay ports that are currently in use), the issue is the building fiber spur cost. One way to address this issue is for the forward-looking carriers to establish joint construction agreements with companies that have rights of way, such as utilities, railroads, bridge and tunnel authorities, and so on. Skepticism is understandable when you realize that the focus of the industry has not been where it needs to be; the focus has been on developing new Layer 2 MAC protocols rather than on lowering the per-building fiber access cost, which is comprised of the spur cost and the in-building equipment cost. "Another important realization in 2001 is that the anticipated opportunities for new data-centric service providers have not materialized, leading to a much slower pace of technology adoption."[10]

Table 1-5 makes another important point while plotting the same data. In this case, technology represents, on average, 12.4 percent of the expense. Therefore, being extremely smart about a carrier's technological baseline would, in the best instance, modify the bottom line by 12.4 percent, or more realistically, 1 to 3 to 5 percent improvement. Transmission, mostly in the form of spur costs, represents on the average 33.6 percent of the total cost. As noted, this is an area that urgently needs attention, but it is important to understand that it is generally the building access component that must be improved upon, not the access backbone or the core backbone. (Table 1-6 shows that except in cases where customers pay for a full gigabit service,

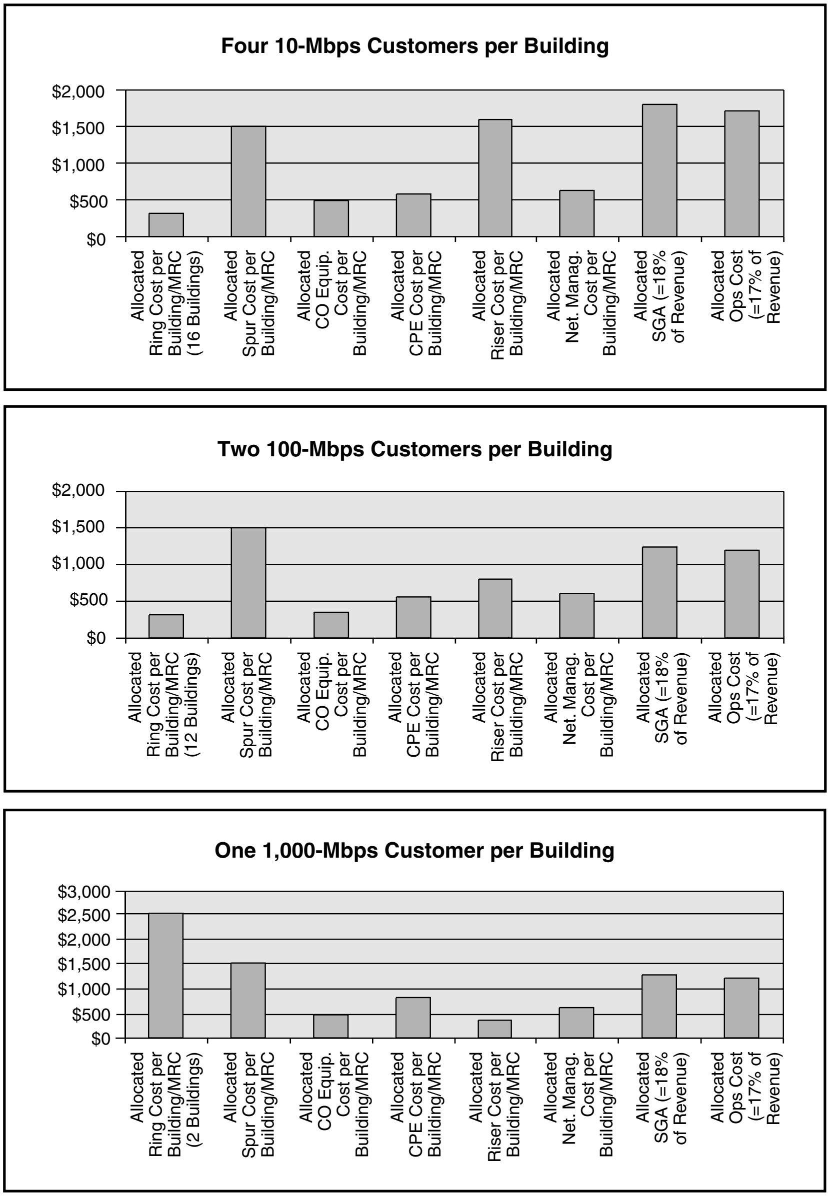


Figure 1-23
Cost elements for broadband services.

the spur can be 65 to 83 percent of the total transmission cost.) Operations (administration, provisioning, engineering, the network operations center [NOC], and so on) is where the major cost factor is—54 percent on the average. Notice that technology costs can be as little as 10.7 percent (two 1,000-Mbps customers in a building on the backbone ring) and as high as 15.5

Table 1-5

Cost Elements by Factors—Metro Applications

	Operations	Technology	Transmission	Total
Two 10-Mbps customers in a building on the backbone ring	$3,150	$909	$1,813	$5,872
	53.6%	15.5%	30.9%	100.0%
Four 10-Mbps customers in a building on the backbone ring	$5,700	$1,056	$1,813	$8,568
	66.5%	12.3%	21.2%	100.0%
Two 100-Mbps customers in a building on the backbone ring	$3,850	$909	$1,917	$6,676
	57.7%	13.6%	28.7%	100.0%
Four 100-Mbps customers in a building on the backbone ring	$7,100	$1,056	$2,333	$10,489
	67.7%	10.1%	22.2%	100.0%
One 1,000-Mbps customer in a building on the backbone ring	$3,450	$1,333	$4,000	$8,783
	39.3%	15.2%	45.5%	100.0%
Two 1,000-Mbps customers in a building on the backbone ring	$6,300	$1,540	$6,500	$14,340
	43.9%	10.7%	45.3%	100.0%
Average	54.0%	12.4%	33.6%	100.0%

Table 1-6

Percentage Spur Cost Compared to Total Transmission Cost—Metro Applications

	Transmission	Spur	% Spur
Two 10-Mbps customers in a building on the backbone ring	$1,813	$1,500	82.8%
Four 10-Mbps customers in a building on the backbone ring	$1,813	$1,500	82.8%
Two 100-Mbps customers in a building on the backbone ring	$1,917	$1,500	78.3%
Four 100-Mbps customers in a building on the backbone ring	$2,333	$1,500	64.3%
One 1,000-Mbps customer in a building on the backbone ring	$4,000	$1,500	37.5%
Two 1,000-Mbps customers in a building on the backbone ring	$6,500	$1,500	23.1%

percent (two 10-Mbps customers in a building on the backbone ring). Notice that transmission costs are as low as 21.2 percent (four 10-Mbps customers in a building on the backbone ring) and as high as 45.5 percent (one 1,000 Mbps customer in a building on the backbone ring).

Finally, a couple of words are necessary about the recent watershed events in the telecommunications and data communications industry that have seen dozens of large companies go out of business and many others undergo major stock market reevaluations. Rather than undertaking a root-cause analysis, which is both very doable and telling, we focus on a hypothetical company with (very simplified) parameters, as shown in Table 1-7.

First, note some generic trends in the revenue. The initial growth is high, but it later levels to the industry average of, say, 30 percent, year over year. In year 6, this company would have a 1 percent market share of the total local telecom/datacom business, which we see at about $85 billion in 2006. A 1 to 3 percent market share is a reasonable number, and it would be extremely difficult and noncredible to postulate a higher penetration rate. Next, note that there are some typical operations expenses (which, for this example, is assumed to include all SGA, operations, equipment leases, and transmission costs/leases). EBITDA is simplified to equate to the gross income, and net income is shown as a specified number in the table. As you can see, the company is not EBITDA positive until the sixth year, the income does not kick in until the same year—in practice, it could have taken even longer—and it is always below the EBITDA value. Because the company is operating at a loss, a debt bottom line is building. In this

Table 1-7

Profile of a Hypothetical, Typical xLEC

Year	YoY Revenue Gross %	Revenue ($ million)	Operations Expenses (All inclusive) ($ million)	EBITDA (gross income) ($ million)	Net Income ($ million)	Cumulative Debt (assumes it's carried, not paid off) ($ million)	Valvation 5 Times Revenue ($ million)	Valuation 15 Times EBITDA ($ million)	Valuation 20 Times Net Income ($ million)	Valuation 20 Times Net Minus Debt ($ million)
1		$6	$20	−$14	$0	−$14	$30	$0	$0	$0
2	8.00	$54	$100	−$46	$0	−$60	$270	$0	$0	$0
3	2.00	$162	$300	−$138	$0	−$198	$810	$0	$0	$0
4	1.00	$324	$500	−$176	$0	−$374	$1,620	$0	$0	$0
5	0.80	$583	$700	−$117	$0	−$491	$2,916	$0	$0	$0
6	0.60	$933	$900	$33	$17	−$491	$4,666	$497	$331	$0
7	0.40	$1,306	$1,100	$206	$103	−$491	$6,532	$3,096	$2,064	$1,573
8	0.30	$1,698	$1,200	$498	$249	−$491	$8,491	$7,474	$4,983	$4,492

simplification, we assume that the debt is carried and is not offset with income from the net income line. In fact, it was the accumulated debt of many LECs and fiber IXCs that put an unhealthy percentage of them in Chapter 7/11 mode in the recent past.

Now comes the punch line of Table 1-7. This table shows a number of ways of calculating company valuation, which drives company stock value (which is assumed to be company valuation divided by the number of issued common shares). As can be seen, many entries in these columns are zeros or nearly zero values. It used to be that the valuation was 10 or 20 times the forward revenues (for example, the acquisition by AT&T of Teleport Communications Group [TCG] for $12 billion, whereas TCG's yearly revenues were around $600 million). That figure has now come down quite a bit to 3 times the revenues for a company with progressive services (maybe 10 times in rare cases), 3 times the revenues for a company with mediocre services, and 0.5 to 1 times the forward revenues for an ISP. However, the market is imposing an even tighter definition of 15 times EBITDA, an even tougher definition of 20 times the net income, or even a definition of 20 times the net income minus debt. As can be seen in the table, under these definitions, valuations come down from the stratosphere quite a bit.[38]

Finally, Table 1-8 represents the approximate financials of a hypothetical ELEC based on anecdotal information reported in the press towards the end of 2001. We assume for the purpose of this hypothetical profile that venture capitalists invested around $300 million into the company.

A majority of ELECs, which at this writing were tied to a single architecture and technology, were relegated by and large to just offering GbE Internet access services rather than having properly deployed a multiarchitecture, multitechnology, multiservice, multi-QoS, and multimedia apparatus. The relevant observation is that, at the time of this writing, traditional ISPs were generally valued by the market at 1.5 times their forward annual revenue (refer to Appendix A for support). ELECs that only offer Internet access service over an Ethernet access/Ethernet architecture platform are destined to be valuated at 1.5 times their annual forward revenue. If and when these ELECs have properly deployed a multiarchitecture, multitechnology, multiservice, multi-QoS, and multimedia apparatus, they can be valuated by the market at 5 to 10 times their forward annual revenue.

The issue of company debt is key. Take as an example ELEC A with $1 million in annualized revenue, $0.2 million annually in SGA, $0.2 million annually in operations cost annually, $0.15 million in (amortized) equipment costs, and $0.15 million annually in recurring facilities costs. ELEC A's bottom line would be $1 million−$0.7 million=$0.3 million. Deducting

Table 1-8

Profile of a Hypothetical, Typical ELEC

Year	YoY Revenue Gross %	Revenue ($ million)	Operations Expenses (All inclusive) ($ million)	EBITDA (gross income) ($ million)	Net Income ($ million)	Cumulative Debt (assumes it's carried, not paid off) ($ million)	Valvation 5 Times Revenue ($ million)	Valuation 15 Times EBITDA ($ million)	Valuation 20 Times Net Income ($ million)	Valuation 20 Times Net Minus Debt ($ million)
1		$1	$14	−$13	$0	−$13	$5	$0	$0	$0
2	8.00	$9	$77	−$68	$0	−$81	$45	$0	$0	$0
3	2.00	$27	$200	−$173	$0	−$254	$135	$0	$0	$0
4	1.00	$54	$300	−$246	$0	−$500	$270	$0	$0	$0
5	0.80	$97	$350	−$253	$0	−$753	$486	$0	$0	$0
6	0.60	$156	$350	−$194	$0	−$947	$778	$0	$0	$0
7	0.40	$218	$350	−$132	$0	−$1,080	$1,089	$0	$0	$0
8	0.30	$283	$350	−$67	$0	−$1,147	$1,415	$0	$0	$0

$0.12 million in taxes (a 40 percent rate), ELEC A has a net of $0.18 million, or an 18 percent net margin. Over 5 years this company would have generated $0.9 million of net income. Table 1-9 shows that a valuation of 5 times the forward revenue has an IRR of the project at −39 percent. A valuation at 0.5 times the forward revenue has an IRR of 23 percent. Notice that a metric that stated 3 times the forward net income ($0.54 million) produced practically the same results.

Now consider ELEC B with $1 million in annualized revenue, $0.2 million annually in SGA, $0.2 million annually in operations costs, $0.15 million annually in (amortized) equipment costs, and $0.15 million annually in recurring facilities costs. Also assume that the company has a $1 million

Table 1-9

Impact of Debt on ELEC's Valuation

ELEC A		**ELEC B**	
Year 0 (valuation)	−$5.00	Year 0 (valuation)	−$5.00
Year 1	$0.18	Year 1	$0.04
Year 2	$0.18	Year 2	$0.04
Year 3	$0.18	Year 3	$0.04
Year 4	$0.18	Year 4	$0.04
Year 5	$0.18	Year 5	$0.04
IRR	−39%	IRR	−58%
Year 0 (valuation)	−$1.00	Year 0 (valuation)	−$0.50
Year 1	$0.18	Year 1	$0.04
Year 2	$0.18	Year 2	$0.04
Year 3	$0.18	Year 3	$0.04
Year 4	$0.18	Year 4	$0.04
Year 5	$0.18	Year 5	$0.04
IRR	−3%	IRR	−25%
Year 0 (valuation)	−$0.50	Year 0 (valuation)	−$0.11
Year 1	$0.18	Year 1	$0.04
Year 2	$0.18	Year 2	$0.04
Year 3	$0.18	Year 3	$0.04
Year 4	$0.18	Year 4	$0.04
Year 5	$0.18	Year 5	$0.04
IRR	23%	IRR	23%

debt, which it is repaying in five years ($0.2 million per year), plus interest of $0.036 million per year (6 percent interest on a $1 million loan). ELEC B's bottom line would be $1 million−$0.936 million=$0.064 million. Deducting $0.025 million in taxes (a 40 percent rate), ELEC B has a net of $0.039 million, or a 3.9 percent net margin. Over 5 years this company would have generated $0.195 million of net income. Table 1-9 shows that a valuation of 5 times the forward revenue has an IRR of the project at −58 percent. A valuation at 0.5 times the forward revenue has an IRR of −25 percent. To achieve a comparable IRR as in the previous case, the valuation here would have to be 0.11 times the forward revenue. Notice that a metric that stated 3 times the forward net income ($0.117 million) produced practically the same results.

In both cases, using the same multiplier of forward net income produced identical results. The problem, though, is that startups generally do not have positive net income for a number of years (typically three to five years); therefore, a metric based on forward revenue (or some more complex formula that takes into account other kinds of value) has to be used. Although there may be a degree of art in valuating a company, such as a startup ELEC or an ISP, there is also a good degree of science, as suggested in the previous example.

1.7 Modeling Parameters

This section collects some modeling data that may be useful for the discussion at hand. The information comes from dozens of sources including New Paradigm Resources Group, Inc., Vertical Systems, CIR, Cisco Systems, NGN Ventures 2001, market research firms, and has been interpreted by the authors.

Table 1-10 provides a gamut of information. Above the black line, we show the entire market for local data services, based on information from New Paradigm Resources Group, Inc. and additional sources and forecasting. Note that the growth rate of the bottom line is 30 percent for 2002 to 2003, 28 percent for 2003 to 2004, 26 percent for 2004 to 2005, and 21 percent for 2005 to 2006. These percentages are revised by New Paradigm Resources Group, Inc. and are much smaller and conservative than the ones New Paradigm Resources Group, Inc. and the industry were using earlier on. Next, based on data from Vertical Systems, NGN Ventures 2001, and additional sources, the data market is allocated across a low end (made up of sub-T1 services), mid end (T1-level services), high end (10-Mbps Ether-

Table 1-10

Market Numbers, Customer Numbers, and Building Numbers Applicable to Broadband

		2002	2003	2004	2005	2006
Low end	Sub-T1	$19,150,957,854	$23,968,276,338	$29,428,291,791	$35,420,523,540	$40,749,023,117
Mid end	T1	$9,575,478,927	$12,614,882,283	$16,303,762,765	$20,656,377,630	$25,014,555,209
High end	T3 (also E)	$2,393,869,732	$3,319,705,864	$4,516,277,774	$6,023,145,540	$7,677,827,889
	OC-3 (FE)	$1,196,934,866	$1,747,213,613	$2,502,092,950	$3,512,550,249	$4,713,179,235
	OC-12	$598,467,433	$919,586,112	$1,386,201,080	$2,048,432,862	$2,893,273,830
Very high end	OC-24 (GbE)	$425,576,841	$688,345,160	$1,092,236,000	$1,698,980,456	$2,525,995,556
	OC-192 (10 GbE)	$212,788,421	$362,286,926	$605,116,898	$990,803,590	$1,550,629,941
	Total market	$33,554,074,074	$43,620,296,296	$55,833,979,259	$70,350,813,867	$85,124,484,779
	Broadband portion	$4,827,637,292	$7,037,137,675	$10,101,924,702	$14,273,912,697	$19,360,906,452
	Connections at $4,500/m, eroded 15% a year	89,401	149,865	247,404	402,015	627,080
	Buildings with fiber (15.5% CAGR)	18,000	20,790	24,012	27,340	32,033
	Buildings per city (30 cities have 50%)	300	347	400	462	534
	Buildings per city (next 60 cities have remaining 50%)	150	173	200	231	267
	Orders per city (30 cities have 50%)	1,490	2,498	4,123	6,700	10,451
	Orders per city (next 60 cities have 50%)	745	1,249	2,062	3,350	5,226

Table 1-10 cont.

Market Numbers, Customer Numbers, and Building Numbers Applicable to Broadband

	2002	2003	2004	2005	2006
Broadband customer/ building (30 top cities)	5	7	10	15	20
Broadband customer/ building (next 60 cities)	5	7	10	15	20
Broadband customer/ building using CIR building statistics	5	7	10	15(#)	20(#)
Broadband customer/ building, first mover (50%)	3	4	5	8(#)	10(#)
Broadband customer/ building, second mover (25%)	1	2	3	4(#)	5(#)
Broadband customer/ building, all other (25%)	1	1	2	3(#)	5(#)

(#) This number may turn out to be high, but we follow the model's formulation.

net, T3, 100-Mbps Ethernet, OC-3, and OC-12), and very high end (GbE and OC-192/10 GbE). Only the high-end and very high-end strata fit our definition of broadband. Figures 1-24 and 1-25 display this information graphically. Supplementary Table 1-11 depicts this information in terms of percentage splits of total revenue over the years.

Below the black line, the table computes the number of broadband connections that correspond to the cumulative number, assuming a typical TLS charge of $4,500. Some TLS services could cost less and some could cost more; $4,500 represents a value averaged over a set of 34 services at a typical metro service provider. This figure is eroded over time at 15 percent per year in the model. Next we show the number of broadband-ready buildings

Figure 1-24
Revenue allocation (2002).

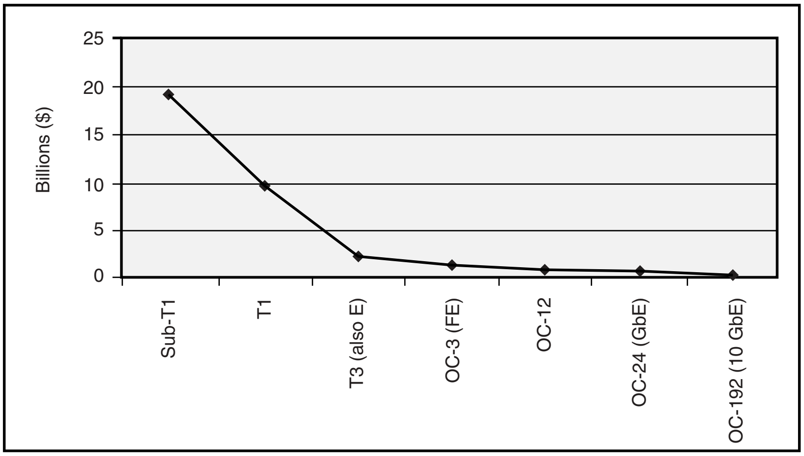


Figure 1-25
Revenue allocation (2006).

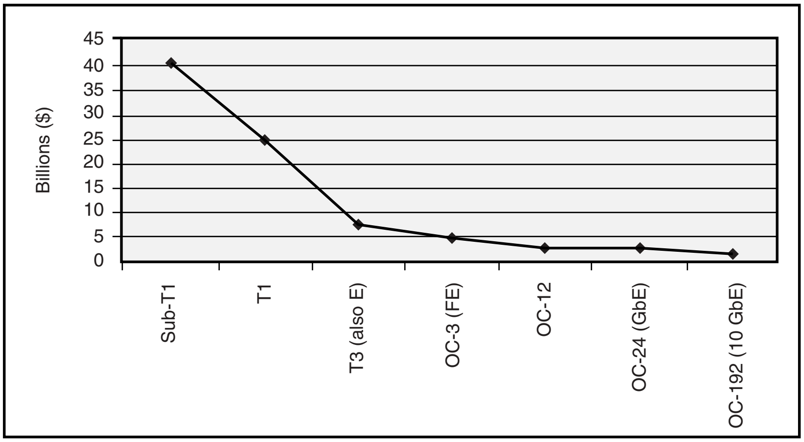


Table 1-11

Percentage Splits of Total Revenue over the Years

		2002	2003	2004	2005	2006
Low end	Sub-T1	57.07%	54.95%	52.71%	50.35%	47.87%
Mid end	T1	28.54%	28.92%	29.20%	29.36%	29.39%
High end	T3 (also E)	7.13%	7.61%	8.09%	8.56%	9.02%
	OC-3 (FE)	3.57%	4.01%	4.48%	4.99%	5.54%
	OC-12	1.78%	2.11%	2.48%	2.91%	3.40%
Very high end	OC-24 (GbE)	1.27%	1.58%	1.96%	2.42%	2.97%
	OC-192 (10 GbE)	0.63%	0.83%	1.08%	1.41%	1.82%

based on data from CIR, Cisco Systems, NGN Ventures 2001, and other sources. Notice that there is a compound annual growth rate (CAGR) of 15.5 percent. We then show the average number of fibered buildings in cities by assuming that the top 30 cities in the United States have half the fibered buildings, and that the next 60 cities have the other half of the fibered buildings. This information, in turn, enables us to heuristically calculate the number of broadband orders for two tiers of cities: the top 30 cities, and then the next 60 cities.

Finally, we derive an average of the number of broadband customers per building, which we believe is one of the most pertinent parameters to establish. For example, we see that in 2002 there are, according to these derivations, 5 broadband customers per building and 11 in 2006. It is assumed that the carrier that has first access to the building will capture 50 percent of the customer base; the second carrier will capture 25 percent of the base; and all other carriers share the remaining 25 percent of the base. If you look at a hypothetical second carrier, it will have access to one broadband customer per building in 2002, two customers in 2003 and 2004, and three customers in 2005 and 2006. Notice that in our derivation, one customer proxies for the equivalent of $4,500 of revenue per month. We could have derived a higher number of customers assuming a lower per-order revenue. However, the revenue per building, which ultimately is the true measure of interest because it dictates and/or drives the architectural design and the cost of the equipment that can justifiably be deployed in a building, would be identical in either case.

Some anecdotal information is already emerging from customers and the real-estate industry that some companies may be looking to move and/or construct low-rise buildings. This would only reinforce the point made that the number of broadband customers per building may be/become relatively low, rather than as high as some people in the ELEC industry are claiming (refer to Table 1-10). This anecdotal information also reinforces the point that besides other key figures of merit, the equipment cost per building must be low; consequently, low(er)-end technology such as next-generation SONET may be better suited in the future for the metro access/metro core than the more expensive DWDM.

ELEC's chapter 11 events of the recent past validate the analysis previously provided.

Appendix A

ELECs with Internet-only service over an Ethernet platform are valuated at 1.5 times the annual forward revenue. This appendix documents how value was established (circa 2001) for an ELEC with Internet-only services, basically an ISP on an Ethernet MAN infrastructure, at approximately

$$\text{value} = 1.5 \times \text{annual revenue}$$

Typically, such ELECs offer 100-Mbps or 1-Gbps access to the Internet for $1,000 or $1,500, respectively.

The analysis is based on 10 publicly traded ISPs, using data from October 25, 2001.

The range of valuations described in the following section is from about 1 to 4 times; 1.5 times is the working number that is tied to the most recent market capitalization numbers.

The evaluation is for generic ISPs. The same value ratio derived in the following section is applied to any kind of ELEC that focuses on providing only Internet services.

The open-press analysis is comprised of four elements, in order of importance, as follows:

1. A (recent) article by ISP-Planet Staff on ISP values per subscriber. Once an average value per subscriber is derived, a revenue figure per subscriber a year is utilized to arrive at the ratio (in fact, three revenue possibilities are used and the one producing the highest valuation is utilized). (Notice that the higher the per-month subscriber revenue is assumed, the lower the value multiplier factor.)
2. A set of relatively recent acquisitions (even going back a few quarters when the value of ISPs may have been higher). Once an average value per subscriber is derived, a revenue figure per subscriber a year is

utilized to arrive at the ratio (in fact, three revenue possibilities are used and an average valuation is utilized).

3. An article by ISP-Planet Staff on how to value an ISP.
4. The price of customer data from MCG Capital Corporation.

A.1 Modeling Item #1

The material that follows is verbatim from ISP-Planet Staff, "Subscriber Values," October 25, 2001, http://www.isp-planet.com/research/subvalues/index.html.

It is buyout time in the ISP world as the status of Excite@Home (NASDAQ: ATHM), Prodigy (NASDAQ: PRGY), and High-Speed Access Corp (NASDAQ: HSAC) are all in flux. Juno and NetZero have merged into United Online (NASDAQ: UNTD). We list public companies only and do not include ISPs that are part of a much larger firm, like MSN, for example. Current subscriber counts are derived from corporate reports and ISP-Planet's ISP Rankings. WSRN.com is the source for market data. Market capitalization data is as of market close on October 24, 2001. Excite@Home (NASDAQ: ATHM) is in Chapter 11, Prodigy (NASDAQ: PRGY) is being bought out, and other ISPs continue to make financial arrangements. RCN (NASDAQ: RCNC) issued bonds, and High-Speed Access Corp. (NASDAQ: HSAC) is selling parts of its business. The merged NetZero and Juno started trading as United Online (NASDAQ: UNTD).

A.1.1 Data

Third-party subscriber value research from MCG Capital Corporation averages the ISPs that MCG bought and sold last year. The revenue per subscriber a month is particularly interesting data. Note, however, that the rate of ISP deals is declining drastically.

Stock Symbol	ISP	Value per Subscriber	Market Cap (millions)	Number of Subscribers
[ELNK]	Earthlink	$372	$1,837	4,900,000
[PRGY]	Prodigy	$141	$464	3,300,000
[LOAX]	Log On America	$80	$2.4	30,000

(*continued*)

Stock Symbol	ISP	Value per Subscriber	Market Cap (millions)	Number of Subscribers
[UNTD]	United Online	$69	$78	884,000
[GEEK]	Internet America	$16	$2.4	147,000
[RCNC]	RCN	$582	$295	507,000
[WGAT]	WorldGate	$229	$51	104,000
[HSAC]	High-Speed Access	$102	$18	176,000
[GOAM]	GoAmerica	$427	$43	100,647
[MTNT]	Motient	$84	$21	250,000
Total		$270	$2,812	10,398,647
Assumed yearly revenue per subscriber	19.95×12	1.129 (ratio of yearly revenue to value)		
Assumed yearly revenue per subscriber	15.95×12	1.412 (ratio of yearly revenue to value)		
Assumed yearly revenue per subscriber	29.95×12	0.752 (ratio of yearly revenue to value)		

A.2 Modeling Item #2

For additional details, see http://www.isp-planet.com/research/subvalues/subvalues_september_2000.html.

ISP Subscriber Sale Value: $428 (Based on 2000 and 1999 Sales)

Date	Buyer/ Seller	Value per Subscriber	Sale Price (millions)	Number of Subscribers
6-00	Earthlink/OneMain.com	$428	$308 (approx.)	720,000
5-00	Prodigy/FlashNet	$275	$55 (approx.)	200,000
3-00	CoreComm/Voyager.net	$1,405	$506.1056 (approx.)	360,000

(*continued*)

ISP Subscriber Sale Value: $428 (Based on 2000 and 1999 Sales)

Date	Buyer/ Seller	Value per Subscriber	Sale Price (millions)	Number of Subscribers
1-00	OneMain/Radiks.net	$536	$7.5	14,000
10-99	Voyager/ComNet, TDI Internet, Choice Dot Net*	(avg.) $499	$12.5	25,000
9-99	Voyager/MichWeb, Internet Connection Services*	(avg.) $41	$1.2	30,000
Assumed yearly revenue per subscriber		19.95×12	1.787 (ratio of yearly revenue to value)	
Assumed yearly revenue per subscriber		15.95×12	2.236 (ratio of yearly revenue to value)	
Assumed yearly revenue per subscriber		29.95×12	1.190 (ratio of yearly revenue to value)	

*Composite figures: Prices and subscriber counts for individual acquisitions were not available.

A.3 Modeling Item #3

The material that follows is cited from Christopher M. Knight, “How to Price Your ISP When It Is Time to Sell,” July 12, 1999, http://www.isp-planet.com/business/knight_valuation.html.

The key to maximizing return on your years of investment and hard work is finding the buyer for whom *your* assets have the highest value. You have made the big decision. After years of building your successful ISP business, it is time to bail out, cash out, get out, and burn rubber down the road into the sunset . . . but how do you put a price tag on your ISP? How much should you ask and how much can you get? For a market to exist, a willing seller must be matched with a willing buyer. It does not matter how much you want if there is no one in the market willing to pay what you’re asking. So the goal is to look for the buyer that can maximize, leverage, utilize, and value *your* ISP most highly.

A.3.1 ISP Industry Idiosyncrasy

Outside of the ISP industry, many nonpublic businesses are valued based on some percentage of annualized sales revenue (typically one to five times annual earnings). Historically, however, the ISP industry appears to prefer to price companies based on a value per subscriber, which is then multiplied by the number of subscribers to arrive at an asking price.

That is the theory. In practice, buyers' contracts often specify that the seller gets paid only for the subscribers that remain with the new ISP after the deal. Typically, if you get more than 80 percent to convert to the new ISP, you are doing well.

Imperfect as it may be, the subscriber value method is a good place to start in determining an asking price. Some of the major factors that help determine a value per subscriber are

- The size of your subscriber base (bigger is better)
- Your average revenue per subscriber
- Your geographic POP situation in relation to the buyers needs
- The percentage of your subscribers paying by credit card or electronic draft (as opposed to printing invoices or terms)
- Your growth rate in terms of your marketing flywheel and its capability to keep bringing in new subscribers after the deal is made
- Your churn ratio (lower is better)
- Affiliate, associate, or computer dealers signed up to hawk your services
- Who will be paying your early circuit- or service-termination expenses, if any

TIP: *For the typical ISP with around 2,500 subscribers, the going rate (July 1999) is between $150 and $400 per subscriber. Your mileage may vary. The larger you are, the more power you have to fetch higher valuations.*

A.3.2 Physical Assets

Note that the previous discussion says nothing about the value of your network, servers, dial-up access switches, vendor agreements, or anything other than customer/subscriber information. Some buyers will ask for these assets to be lumped into the deal, but you could sell them separately—or, if it is part of the whole deal, you could sell them itemized, so you know what you are getting for them.

A.3.3 Those Pesky Details

Finally, the following are a number of items that many ISPs fail to think about or plan for when they sell out:

- Getting paid, post deal, for additional subscribers that keep coming from the marketing flywheel you spent years building
- A low conversion rate to the new ISP because the transition was not handled carefully or smoothly
- Early T1 or circuit termination, or lease shutdown charges (these can add up to tens of thousands of dollars)
- Yellow page and other marketing costs under contract for the remaining year on which you forgot to get your new buyer to pick up costs
- What to do if the ISP you sell to botches up completely and your reputation gets dragged through the mud because of their incompetence

TIP: *Seek out the ISPs who are seeking to aggregate a couple dozen regional/local ISPs to take them public in the near future. These deals can bring you much more than the standard ISP subscriber-based buyout—if you do not mind the wait and the risk.*

One other resource you may want to check out is ISP-Planet's Subscriber Values table, which lists valuations per subscriber of the largest publicly traded ISPs.

A.4 Modeling Item #4

The material that follows is from Market Data from MCG Capital Corporation (as reported at http://www.isp-planet.com/research/subvalues/subvalues_january_2001.html#mcg).

By Number of Subscribers			
Subscribers	**Less than 5,000 Average**	**5,000 to 20,000 Average**	**Greater than 20,000 Average**
Price per $ of revenue	$1.98	$1.9	$4.75
Price per subscriber	$873	$463	$3,713
Revenue/sub/month	$52.23	$26.63	$24.22
Last 12 Months (6/99–6/00)			
	Average	**Median**	
Price per $ of revenue	$3.6	$2.34	
Price per subscriber	$1,199	$489	
Revenue/sub/month	$35.27	$20.83	
Residential ISPs (less than $30 rev. per sub.)			
	Average	**Median**	
Price per $ of revenue	$3.06	$2.25	
Subscriber multiple	$563	$478	
Revenue/sub/month	$20.44	$19.4	
Business ISPs			
	Average	**Median**	
Price per $ of revenue	$4.00	$2.35	
Subscriber multiple	$3,105	$667	
Revenue/sub/month	$3,960	$166.67	

(continued)

Private Market Deals	Average	Median
Price per $ of revenue	$3.11	$2.27
Price per subscriber	$719	$476
Revenue/sub/month	$34.96	$21

End Notes

1. We have previously called these access networks neighborhood area networks (NANs).
2. This is derived by assuming that 10 percent of the traffic is destined for the outside and that an inside-to-outside concentration of 10 to 1 is suitable.
3. D. Minoli, P. Johnson, and E. Minoli, *Ethernet-Based Metro Area Networks* (New York: McGraw-Hill, 2002).
4. White Paper, "An Optical Ethernet Business Case for Incumbent Local Exchange Carriers—A Nortel Networks Business Case," Publication 56049.25-09-01, September 2001.
5. Jon Cordova, "A Trend That Won't End," in "The Analyst's Corner," *Primedia Business Magazine and Media* (September 28, 2001), http://industryclick.com/newsarticle.asp?Newsarticleid=236185&SiteID=3.
6. For example, " . . . carriers have $17 billion invested in ATM infrastructure with no replacement in sight," Vint Cerf, *The ATM & IP Report* (April/May 2001): 1.
7. Figure adopted and extended from T. Day, "Widely Tunable Lasers—Tackling the Challenges of Today's Networks," www.newfocus.com.
8. We assume that the price to the customer has the form $P = (1+p)\times C$, where p is a profit factor ($0.1 \le p \le 0.6$) and C is cost.
9. "Looking back on 2000, we recognize that much of what passed for business planning was wishful thinking. Last year, five leading service providers . . . invested $60B in fiber-optic assets. Their investment has been squandered, because their market capitalization dropped from

$70B to $10B." Excerpted from J. McQuillan, "Broadband Networking at the Crossroad," NGN 2001 Promotional Materials, October 11, 2001.

10. J. McQuillan, "Broadband Networking at the Crossroad," NGN 2001 Promotional Materials, October 11, 2001.

11. See M. Merard and S. Lumetta, "Architectural Issues for Robust Optical Access," *IEEE Communications Magazine* (July 2001): 116 ff.

12. Although Ethernet-based services offered by ELECs are now receiving a lot of attention, the senior author of this book described and advocated this service in 1987 in the Bellcore/Telcordia Special Report SR-NPL-000790, December 1987, "A Collection of Potential Network-Based Data Services," which listed, in addition to approximately 150 forward-looking data services that could be offered by the ILECs, a "CO-based LAN service," now known as TLS. However, it must be said that the idea of a carrier-supported LAN extension to a city level was not such an esoteric concept (as it is made out to be today). Bellcore's SMDS and the IEEE 802.6 MAN standard were mid- to late-1980s carrier-grade broadband services according to our current definition. Under the senior author's tenure at TCG, the carrier deployed nationwide several hundred high-end (10- and 100-Mbps) TLS applications (customers) during the mid-1990s using TDM and ATM techniques. The ELECs only discovered the TLS opportunities recently, in 1998 and 1999.

13. At an average of four ADMs per SONET ring and $80,000 per ADM, this would equate to an original install cost of $32 billion.

14. The GbE backbones use full-duplex building-to-building links; these links are back-to-back connected to form a ring architecture. The CSMA/CD discipline is not used (which would have been a STDM technique), but the segment is dedicated to a single source instead (one side of the full-duplex segment dedicated to the uplink, whereas the other side of the full-duplex link is dedicated to the downlink). Therefore, the STDM technique is trivialized.

15. Nortel's White Paper, "A Layered Architecture for Metro Optical Networks," 56018.25/08-00, August 2000.

16. Vint Cerf, *The ATM & IP Report* (April/May 2001): 1.

17. The publication date for this report is September 21, 2001.

18. Gartner Group, February 2001, as described in http://grouper.ieee.org/groups/802/3/efm/public/may01/kelly_1_0501.pdf.

19. IEEE 802.3ae quoting Electronicast.

20. KMI Corp, October 15, 2001, www.kmicorp/fiberoptics_market_studies/broadband. According to the same source, by the end of 2001, the FTTH market in the United States was estimated at 89,000 homes, and the FTTC market was estimated at 915,000 homes.

21. www.cir-inc.com/index.cfm?loc=news/view&id=85 notes that "CIR predicts that the metro wholesale wavelength services market could reach almost $900 million (U.S.) in five years' time growing from a mere $30 million (U.S.) in 2001 because of the increasingly attractive value of wavelength services over that of dark fiber as well as the gradual penetration of metro DWDM systems into carrier networks. CIR states that 30-fold growth might be realistic, but cautions against the market rhetoric that some analysts have absurdly engaged in by quoting future estimates of $4-billion metro wholesale service markets."

22. www.cir-inc.com/index.cfm?loc=news/view&id=85, October 18, 2001.

23. Cisco sources; CIR Inc.; et al. NGN Ventures 2001.

24. The first volume in this set (D. Minoli, P. Johnson, and E. Minoli, *Ethernet-Based Metro Area Networks* [New York: McGraw-Hill, 2002]) provides additional information on the major market opportunity that exists.

25. "IP Growth Revisited," *The ATM & IP Report,* 8, no. 3 (July/August 2001): 1 ff.

26. L. Roberts, Materials at www.caspiannetworks.com.

27. Another technical effort is to build systems that are unrepeated over long distances, but these advances are less applicable to the metro access/metro core environment than in the long-haul/international environment.

28. N. Cole, J. Hawkins, M. Green, R. Sharma, and K. Vasani, "Resilient Packet Rings for Metro Networks," a report by the Resiliant Packet Ring (RPR) Alliance, 2001.

29. Agilent Press Release, Palo Alto, Calif., May 17, 2001.

30. In this context, we mean *metro core bandwidth*, although you could also include here *long-haul bandwidth*.

31. Because a packet service is bandwidth efficient, and therefore theoretically cheaper, the statistical nature of the service may/should occasionally show some warts, which is part of the bargain. However, to hide these aspects of the service from the user, carriers have tended to overengineer the networks, thereby adding more bandwidth than necessary, undoing the supposed advantages of bandwidth parsimony.

32. This signaling protocol enables the dynamic establishment of optical connections in an ION.

33. G. Lee, "SONET: A Viable Solution to the Metro Access Network Bottleneck?" Opticalbiznet.com, June 2001.

34. Light Reading/Optical Oracle at http://www.lightreading.com/document.asp?site=lightreading&doc_id=8340.

35. Roy Bynum, "The Reality of the Transmission Infrastructure," IEEE 802.17, January 2001.

36. In practical terms, not the entire 10^{14}-Hz spectrum can be used because of the attenuation characteristics of the fiber. Perhaps 200 nm (for example, 1,300 to 1,400 nm and 1,500 to 1,600 nm) can be used. Noting that $f = c/\lambda$, 1,300 nm equates to a frequency of 2.30×10^{14} Hz and 1,400 nm equates to a frequency of 2.14×10^{14} Hz; the bandwidth is then 1.64×10^{13}, and with 1 to 10 bits per Hz, you obtain a digital throughput of 16 Tbps to 160 Tbps. 1,500 nm equates to a frequency of 2.00×10^{14} Hz and 1,600 nm equates to a frequency of 1.87×10^{14} Hz; the bandwidth is then 1.25×10^{13}, and with 1 to 10 bits per Hz, you obtain a digital throughput of 12 Tbps to 125 Tbps. Hence, the total theoretical digital throughput is around 300 Tbps or 100,000 more than a 2.4-Gbps-based system. Another limiting issue is amplifier performance. Although no one has achieved these kinds of theoretical results, it is clear that the fiber can carry much more information than is achievable on a standard pre-DWDM system.

37. http://www.lightreading.com/document.asp?site=lightreading&doc_id=3006&page_number=8.

38. Perhaps the first analytical view of the valuation of some ELECs came when Cogent Communications Group filed papers with the Securities and Exchange Commission (SEC) in October 2001 regarding the acquisition of Allied Riser Communications. *The Washington Post*, citing the SEC filing, said that the company had raised $62 million through a preferred-stock offering and increased its line of credit from Cisco from $310 million to $409 million. Reportedly, the company billed revenue for the first half of 2001 was only $90,000.

CHAPTER 2

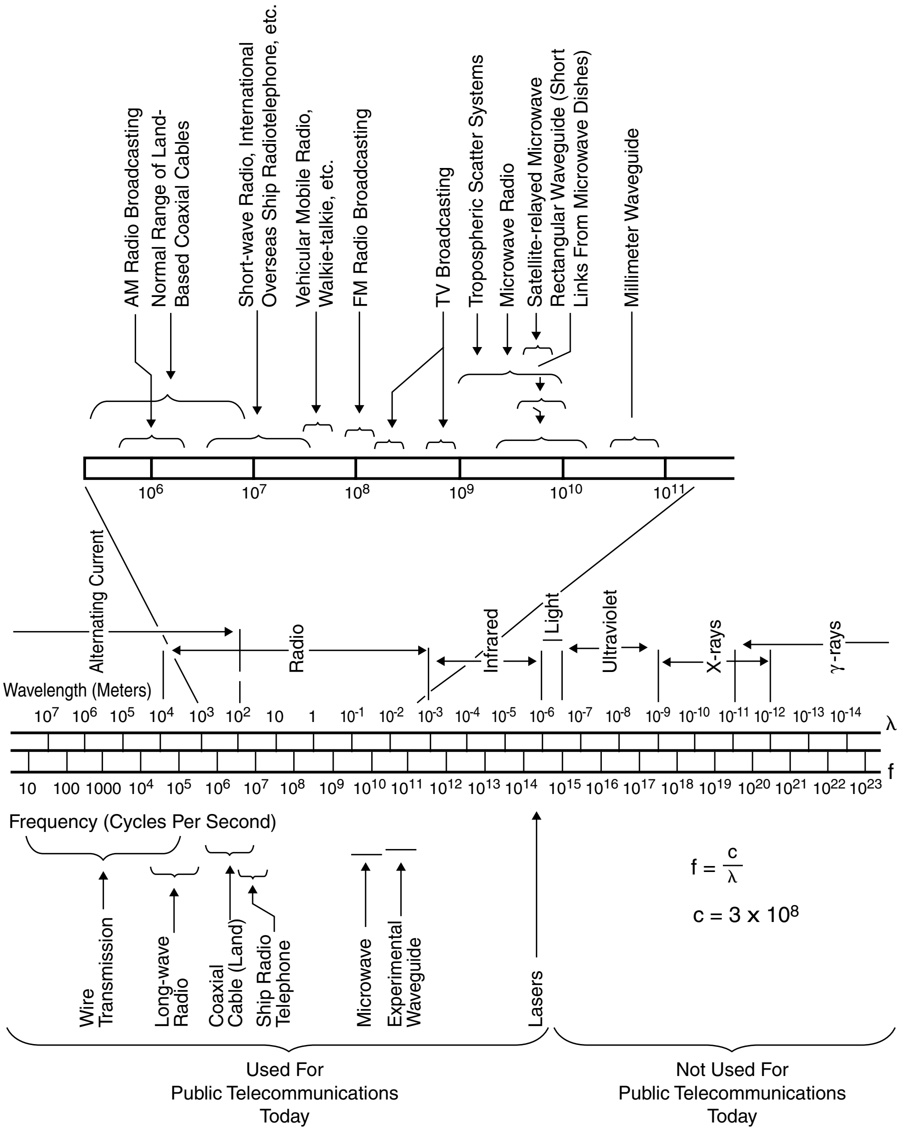
Optics Primer

This chapter provides a short primer on optical technologies and lays the groundwork for the rest of the book. Commercial applications of optical fiber technology for telecommunications purposes were first introduced in the field in 1977. This original system supported a DS3 link that was 2 km (1.5 miles) long. Today 10 Gbps can be supported on a single beam of light[1,2] (and work that aims at supporting 40 Gbps is underway). Multiplexing systems enable the carriage of several dozen beams onto a single fiber or fiber pair. Although unrepeated distances have increased significantly over the years, the use of repeaters creates virtually unlimited distances. Figure 2-1 positions optical communication in the electromagnetic spectrum.

The basic components of a fiber-optic system are a transmitter (laser diode [LD] or light-emitting diode [LED]), a fiber-optic waveguide (single-mode or multimode fiber), and a receiver (photodetector diode), as shown in Figure 2-2. Most commercially deployed systems today are unidirectional (that is, one fiber is used to transmit and another fiber is used to receive). Typically, termination-multiplexing equipment is at the end of the link to support tributary lower-speed signals. This transmission/multiplexing equipment generates electrical signals that the fiber-optic transmitter converts to optical signals for transmission over the optical fiber. The far-end receiver converts the optical signals back to electrical signals for distribution to the various ancillary devices.

When originally introduced, fiber-optic systems were implemented for long-distance applications based on very traditional cost-of-transmission-versus-cost-of-multiplexing economic trade-off considerations. As the costs of the fiber media and multiplexers decreased over time, Local Exchange Carriers (LECs) identified cost-effective applications for fiber-optic systems in the metropolitan (metro core) and loop feeder or distribution environment (metro access). However, the economic considerations for the deployment of loop fiber have up to now mostly been based on supporting DS0 voice applications over traditional digital loop carrier systems operating typically at an aggregated DS3 rate, rather than on supporting data applications at some higher data rate. These traditional applications have covered the feeder portion of the loop plant, connecting the central office (CO)

Figure 2-1
Positioning of optical communication in the electromagnetic spectrum.



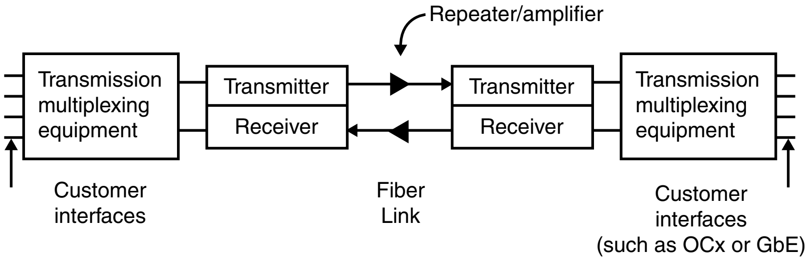


Figure 2-2
Fiber-optic system.

to the remote telco node. As another application, the cable TV industry has replaced in recent years some major portions of their distribution networks with a Fiber to the Node (FTTN) configuration. The economics of deploying fiber to support mid-range or broadband data applications to the home or office are just now receiving widespread attention. The bulk of this chapter focuses, however, on technology rather than on economics and cost-optimized network design.

2.1 Optical Fibers

Optical fibers used in carrier applications are made predominantly from silica (SiO_2). Pure silica has a low optical signal loss in the infrared region of the electromagnetic/electro-optical spectrum. Infrared ranges from 0.80 to 1.60 μm in terms of signal wavelength (800 to 1,600 nanometers).

Refraction is one of the most important properties of interest in optical communication. The difference in the index of refraction gives the optical fiber its transmission (waveguide) properties. Refraction occurs because light moves faster through some materials than through others. The index of refraction of either the core or the cladding (or both) is changed from the values of pure silica by the incorporation of small amounts of dopant materials, such as germanium and phosphorus. The index of refraction is defined in reference to two media (see Figure 2-3) and represents the ratio of the speeds of light propagation in the two media. In physics tables, the value is expressed in reference to a vacuum. The index of refraction is also dependent on the wavelength, as shown in Figure 2-4.

Optical fiber is constructed with a cylindrical core of a higher index of refraction than its cladding material. Figure 2-5 illustrates the fiber core composition of transparent glass with a relatively high index of refraction.

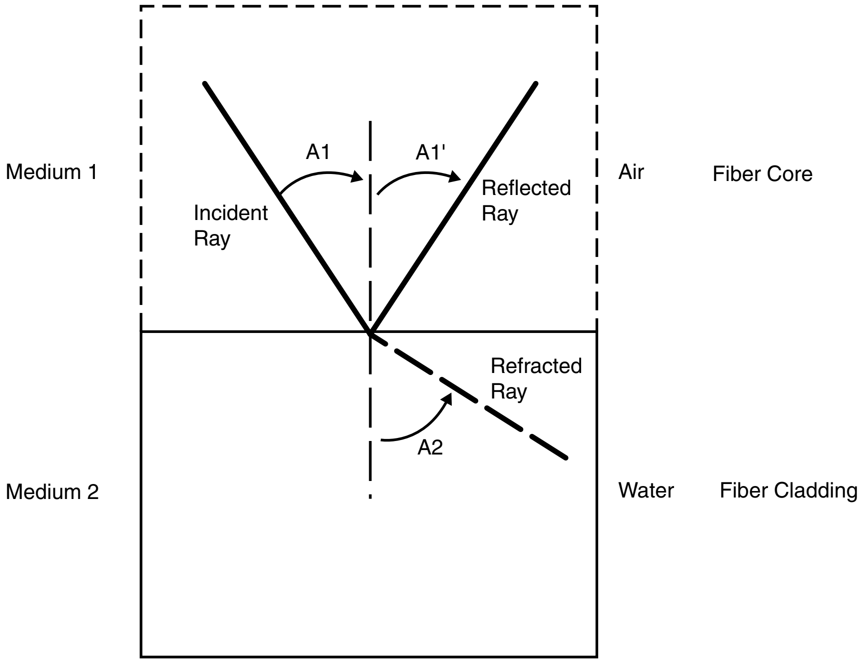


Figure 2-3
Definition of index or refraction.

Notes

$A_1 = A_1'$

$\frac{\sin A_1}{\sin A_2} = n_{21}$ = Index of refraction of medium 2 with respect to medium 1

or, by further derivation

$n_{21} = \frac{v_1}{v_2}$, which is the ratio of the speeds of light in the two media

In turn, the cladding is enveloped by a protective outer cable material, which plays no role in light propagation. The generic specifications for optical fiber and fiber cables for metro and long-haul applications cover the specific requirements for optical, geometric, mechanical, cable-jacket, environmental, and electrical-protection features.

Fiber acts as a waveguide for the signal. Figure 2-6 shows how light travels in a fiber. Because there is a difference of speed in the two media (core and cladding), the light is reflected and refracted at the boundary; the goal is to achieve total internal reflection, so that the light beam is projected forward. This occurs when the angle of refraction is 90 degrees. Because, by

Index of refraction of one medium with respect to another medium generally varies with the wavelength of the incident beam/ray

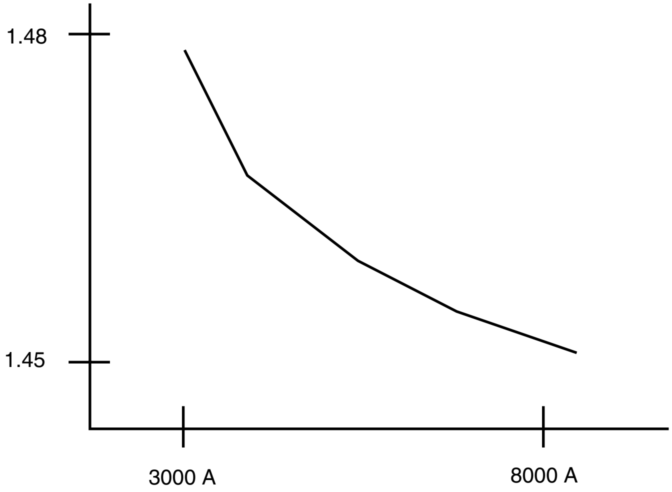


Figure 2-4
Index of refraction's dependence on wavelength.

Some indices of refraction:

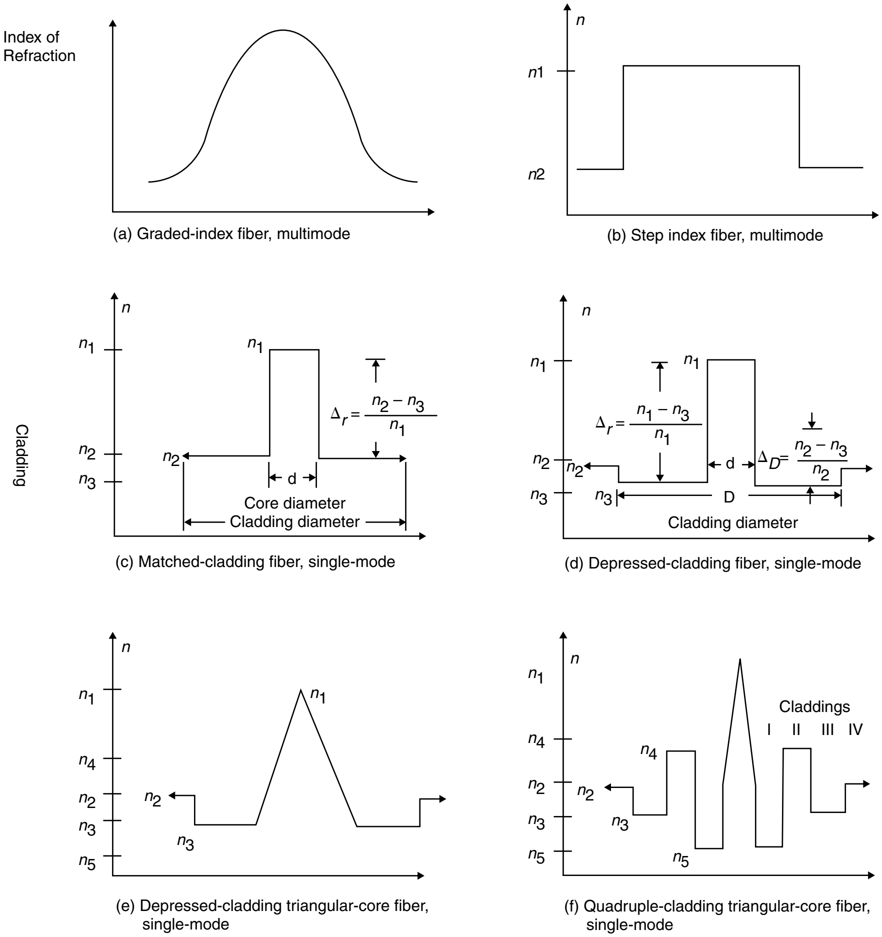
- with respect to vacuum
- At λ = 5890 A

Water	1.33
Air	1.00033
Dense glass	1.66

design, the index of refraction changes abruptly at the core/cladding boundary, the resulting fiber is called *step-index* fiber. There is also a fiber type that has a gradual change in the index across the cross-section of the fiber; this is called *graded-index* fiber.

An optical fiber is manufactured by first fabricating a preform and then performing a process called drawing. The fabrication of the preform uses one of three standard processes. Lucent uses modified chemical vapor deposition (MCVD), which is also called inside chemical vapor deposition (ICVD). Corning uses outside chemical vapor deposition (OCVD), and NTT suppliers use a vapor axial deposition (VAD) process. The fiber is drawn by lowering the preform rod into an oven at a controlled rate; eventually, the preform reaches a high enough temperature to soften, enabling the fiber to be pulled to a specified diameter. The diameter depends on the rate of rota-

Figure 2-5
Fiber refractive index profiles.



tion of the take-up drum and the rate of feed of the preform into the oven (see Figure 2-7).

To form cables, individual fiber strands are placed into slots, tubes, bundles, or ribbons. Several layers of protection may be used in the cable depending on the application. Four major cable structures exist: the slotted core or channel cable, the ribbon cable, the loose tube cable, and the tightly bound cable. In the channel or the loose tube design, one or more fibers are placed in an insulating channel or buffer tube. In the ribbon design, several fibers are combined into a flat ribbon.

Two types of silica-based, optical step-index fiber exist, as shown in Figure 2-8: multimode and single mode. The transmission properties that are

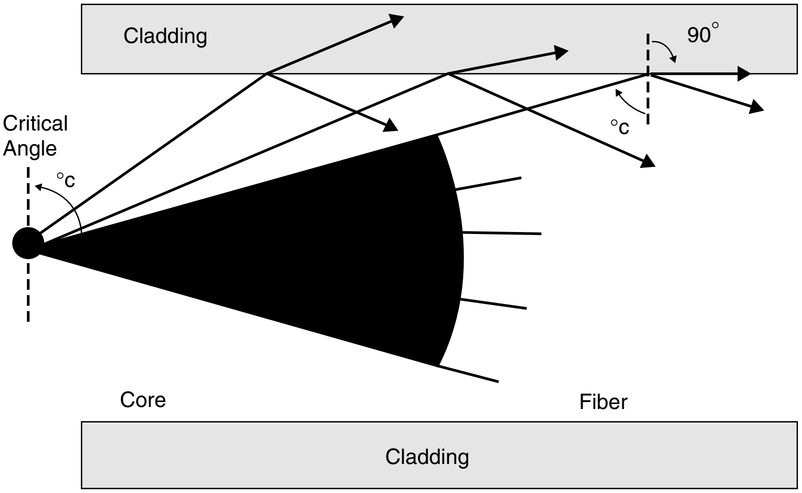


Figure 2-6 How light travels in a fiber.

- Angle at which the refracted ray points along the surface (i.e. angle of refraction = 90 degrees)

- For angles of incidence larger than this critical angle, no refracted ray exists.

  Phenomenon called total internal reflection

- $\Theta_C = \sin^{-1} (n_2/n_1)$

- For air/glass, this is 41.8 degrees.

- Important result: light can be piped from one point to another with little loss.

Note: Core/cladding not shown to scale.

characteristic of the two types of optical fiber arise from the differing core diameters and refraction-index profiles. Figure 2-8 depicts the typical dimensions and refraction-index profiles, showing the cross-section of the cable along the diameter. Note that both optical fibers have an outer diameter of 125 μm, but the core size is different.

Multimode fibers have various ratios of core-to-cladding diameter. The most common ratio is 62.5:125 μm. (Multimode fibers are more typical of local area networks [LANs]; some intraCO wiring is also multimode, but single-mode fiber remains prevalent even in these applications.) Numerical aperture (NA) is a measure of fiber's capability to accept light. Typical fiber with a 50-μm core may transmit several hundred modes. The NA of a mul-

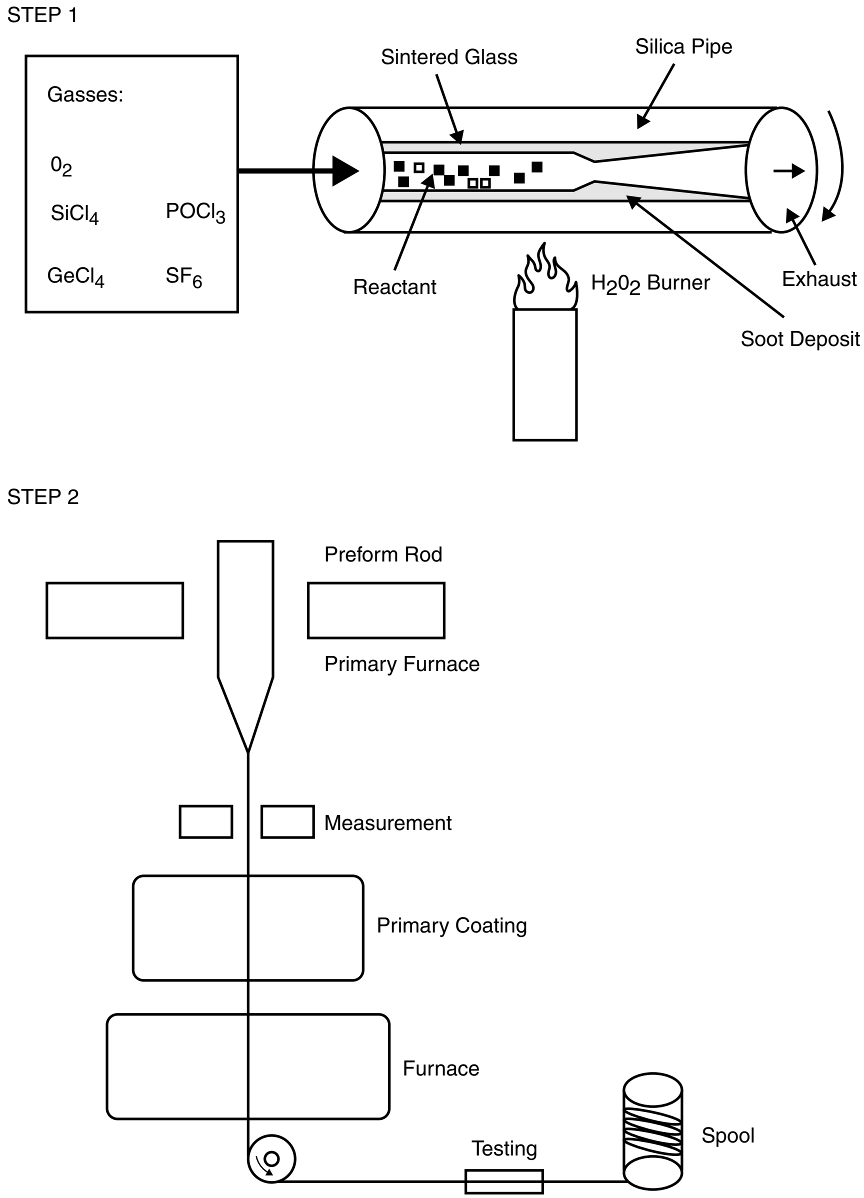
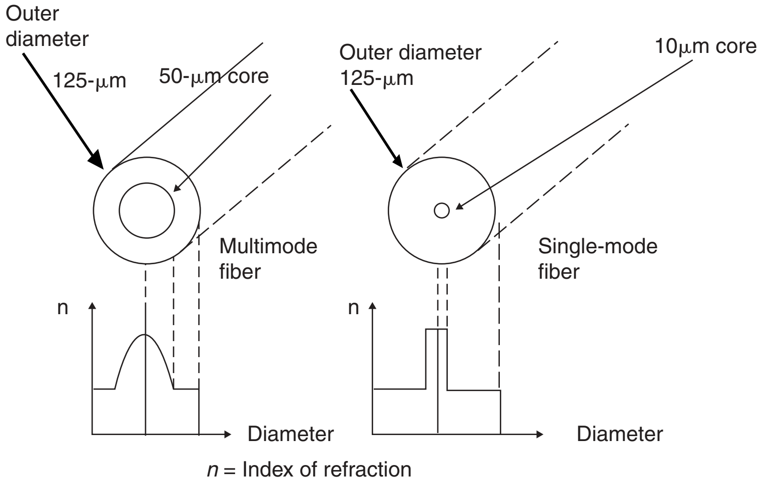


Figure 2-7 Basic fabrication technique.

timode fiber is defined as the sine of the maximum angle of acceptance for coupling optical power into the fiber:

$$\text{NA} = \sin a,\ 2a = \text{Full acceptance angle}$$

Figure 2-8 Characteristics of fiber strands.



Nominal NAs are in the range of 0.20 to 0.29 μm. For an optical fiber in which the index of refraction decreases uniformly and monotonically from n_1 (core refractive index) to n_2 (cladding refractive index), the theoretical NA is given by the following:

$$\mathrm{NA} = \sin a = \sqrt{{n_1}^2 - {n_2}^2}$$

When the core is made smaller, there are fewer modes that carry energy. If the core is made small enough, but not too small, then the fiber only carries one mode. When the core dimension is only a few times larger than the wavelength of the signal to be transmitted, only one of the originating rays (modes) will be transmitted; this results in low loss (the other rays are attenuated in the proximity of the source). For wavelengths of 1.3 to 1.55 μm, the core needs to be less than 10 μm in diameter for single-mode transmission properties to hold. The problem of modal dispersion that affects multimode fiber then disappears. With a single-mode fiber, the information capacity is much higher; however, because the core is so small, more expensive light sources and connectors are required.

Currently, many forward-looking providers install a fiber infrastructure based on cables with a very high fiber count in the order of 500 or even 1,000. Following different service demands, these cables contain different fiber types such as standard single-mode fiber according to International Telecommunication Union Standardization Sector (ITU-T) G.652, nonzero dispersion-shifted single-mode fiber (discussed in the following section) according to ITU-T G.655, or fibers without OH-peak (such as Lucent All-Wave). The main transmission fiber types are standard fiber (nondispersion-shifted fiber [NDSF], such as Corning SMF-28), dispersion-shifted fiber (DSF), and the various types of nonzero dispersion-shifted fiber (NZDSF) (such as Corning LEAF or LS and Lucent Truewave). New transmission fiber types with lower loss in the 1.4-micron region are beginning to emerge together with fibers specialized for the metro space. In this space, however, NDSF is still the most widely used fiber. Corning SMF-28e fiber has recently become available.[3] SMF-28e fiber supports low water peak performance across the entire operating wavelength range, including a specified attenuation at 1,383 nm of less than or equal to 0.31 dB/km with a hydrogen-induced attenuation increase of less than or equal to 0.01 dB/km. This fiber belongs to the family of extended wavelength band (EWB) fibers, as shown in Figure 2-9[4] (see also Table 2-1). SMF-28e fiber is an ITU G.652.C fiber and is compliant with all applicable industry standards. The additional operating wavelength potential opened by low attenuation in the water peak region combined with the emergence of systems designed to operate over those wavelengths translates into more options for network designers. They may choose to transmit more channels on a single fiber to

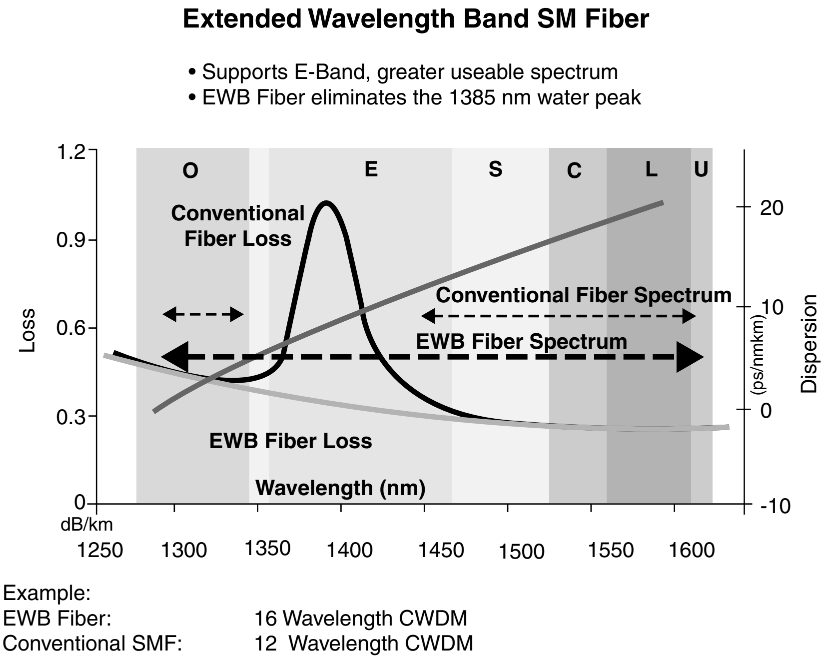


Figure 2-9
EWB single-mode fiber.

Table 2-1

Standards for Basic Fiber Types

	Standard Single-Mode Fiber	**EWB Single-Mode Fiber**
IEC (category) 6073-2-50	B1.1	B1.3 (EWB)
ITU G652	G.652.B	G.652.C (EWB)
TIA 492CAAB (Class)	IVa	IVa (dispersion unshifted with low water peak)

increase channel spacing to facilitate coarse wavelength division multiplexing (CWDM) or reserve additional wavelengths for future use.

2.1.1 Signal Degradation

The fiber medium degrades signal quality, as is the case in a copper medium or any other medium, although degradation is much less than that of copper. This greatly reduced attenuation is the reason why optical fiber can carry a signal over a longer distance or at a much higher bandwidth than copper. Two general kinds of degradation are shown in Figure 2-10:

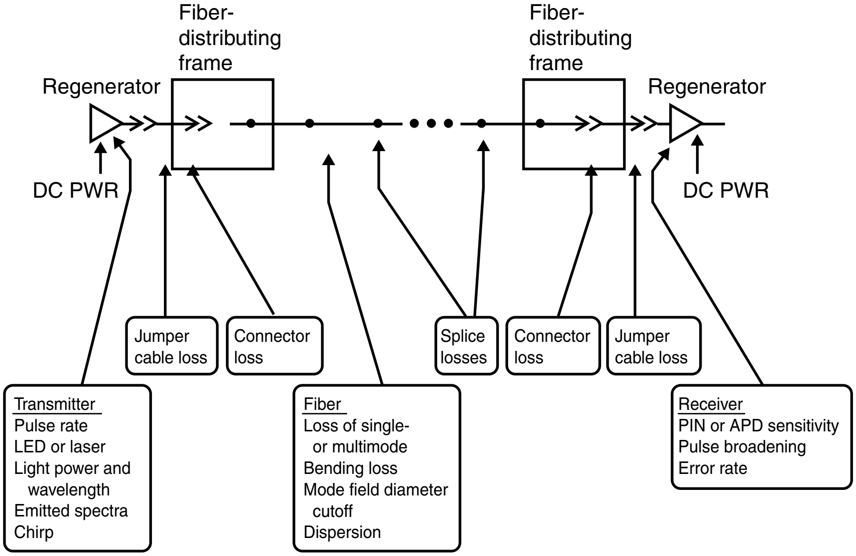
- Attenuation of a signal or waveform
- Distortion of its shape

Attenuation is a well-understood phenomenon. It is the loss of light energy as it travels along the waveguide, that is, along the fiber strand. Energy is lost due to impurities, imperfections, and physical phenomena. Figure 2-11 traces the improvement of attenuation figures of merit over the years.

2.1.2 Dispersion Phenomena

What are the distortion issues? Light not only becomes weaker when it goes through a long segment of fiber, but individual pulses of light in a digital signal may also become broadened or blurred, so that they overlap. Both effects make a signal more difficult to decipher at the remote end; in turn,

Figure 2-10
Losses and design issues for fiber links.



this results in a higher bit error rate (BER). The principle cause of waveform distortion is dispersion. Optical fibers are assigned dispersion ratings expressed either in bandwidth in megahertz per kilometer or in time spread in picoseconds per kilometer. In the latter case, the dispersion of a fiber is measured in picoseconds of pulse spreading per kilometer of fiber per nanometer of source spectral width (ps/km/nm). Bandwidth is limited by the amount of pulse spreading and the response of source and detector.

The total dispersion (pulse spreading or broadening) and the related fiber bandwidth are characterized by two key effects: modal dispersion (also called intermode or multimode dispersion) for multimode transmission systems and chromatic dispersion (also called intramode or spectral dispersion) for single-mode transmission systems.

When a very short pulse of light is injected into an optical fiber, the optical energy does not all reach the far end at the same time. This causes the exit pulse to be broadened or dispersed. Intramodal distortion is manifested in two ways:

1. As material dispersion resulting from slightly different indices of refraction that fiber presents to different wavelengths of light
2. Waveguide dispersion resulting from a longer path taken by longer wavelengths that reflect at more oblique angles with cladding

Figure 2-11 Some fiber optics history.

The laser was invited in the 1960s, and the concept was refined in the next few years.

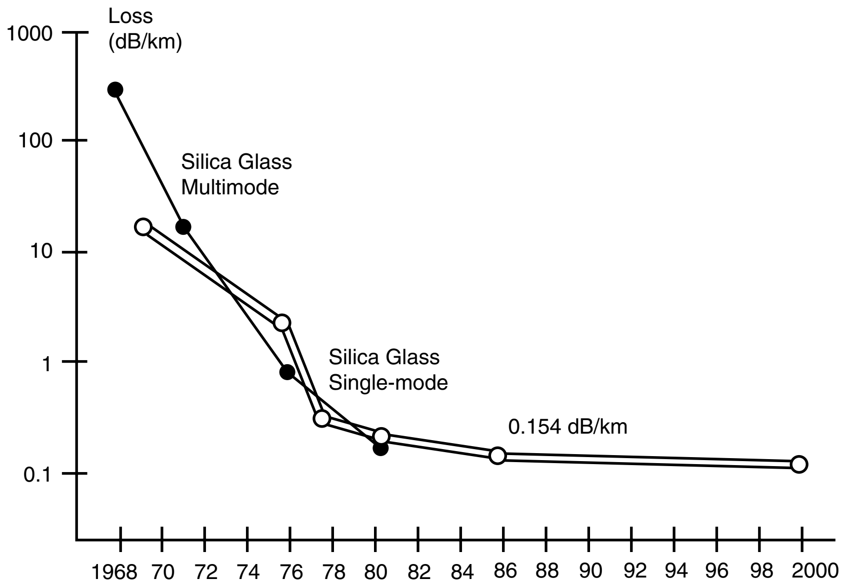
Transmission through optical fiber was proposed in the literature in 1966.

Low-loss optical fiber was announced in 1970.

It was silica-based and it had an attenuation of 20 dB/km.

By 1986, the attenuation had been reduced to 0.154 dB/km.

This is close to the theoretical minimum for silica, this being 0.13 dB/km at 1.55 μm.



Progress of optical fiber in terms of loss, over the years

2.1.3 Intermodal Delay Distortion

In simple step-index multimode fiber (which, as noted, has a large uniform core), rays of light propagate through the core at different angles. Rays that go straight through have a shorter path than those with large angles, which bounce off the inside wall of the core; hence, straight rays traverse the fiber faster. Step-index multimode fiber enables the transmission of any light ray that enters optical fiber within an angle of acceptance. High-order modes travel over a longer distance than low-order modes, and therefore they arrive out of phase at any given point along the fiber. This causes intermodal delay distortion and thus limits the bandwidth capacity of the fiber. Low-order modes, such as those launched into fiber at small angles with respect to the axis, travel a shorter ray distance to reach a given point along the fiber than the distance traveled by high-order modes launched at large angles. Intermodal distortion constricts the bandwidth of the cable because waves that start at the same instant arrive out of phase at the remote end, causing distortion.

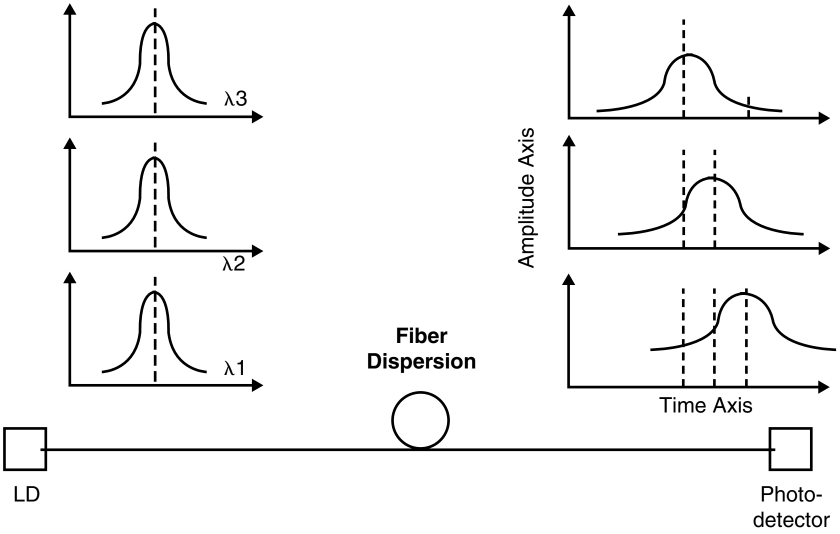
The problem can be reduced by using a graded-index fiber; however, this type of fiber has relatively little if any commercial deployment. With this kind of fiber, high-angle rays, with a longer physical distance to travel, spend more time in regions of the silica/glass with a lower index of refraction where the speed of light is faster. This compensating effect gives all rays almost the same transit time; the result is that graded-index fiber has a much higher information capacity than step-index fiber. The compensation is never perfect, and a practical graded-index multimode optical fiber still cannot carry large data rates more than a few kilometers before distortion makes data unintelligible. Furthermore, it is more difficult to manufacture this type of fiber.

2.1.4 Chromatic Dispersion

Even in a single-mode fiber, all energy in a pulse does not reach the remote end at the same time. This phenomenon is called chromatic dispersion. Chromatic dispersion arises because, although there is only one mode, the light pulse is usually composed of a small spectrum of wavelengths, and different wavelengths travel at different speeds in fiber. Chromatic dispersion is defined as the slope of the transit time versus the wavelength curve for a given fiber substance.

Chromatic dispersion limits the bandwidth of single-mode fibers. It results from the wavelength dependence on the velocity at which the optical signal travels, as shown in Figure 2-4. Optical rays with different wavelengths travel along fiber at different velocities in different media. Because the optical sources (LDs) used produce signals over a range of wavelengths (for example, from LD chirping) and because each wavelength component of the pulse travels at a slightly different velocity, the pulse spreads in time as it travels down the fiber. In the visible spectrum of light, material dispersion causes longer wavelengths to travel faster than shorter ones (see Figure 2-12). However, in the near-infrared region around wavelengths of 1.1

Figure 2-12 Temporal distortion due to chromatic dispersion.

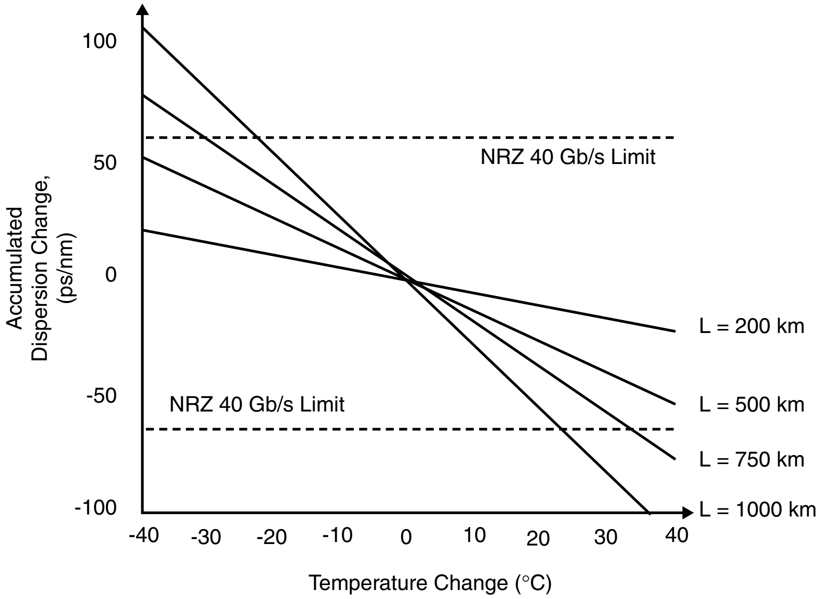


to 1.3 microns, the opposite beings to happen—longer wavelengths begin to travel slower. There is also a temperature dependence in chromatic dispersion, as seen in Figure 2-13. Although this dependence can generally be ignored for speeds up to 10 Gbps, it becomes an issue for higher speeds: an example would be OC768/40 Gbps (its impact is 16 times more severe at 40 Gbps than at 10 Gbps).

The effect of chromatic dispersion can be reduced in three ways:

1. Use a light source with only one wavelength. This idea, however, is somewhat theoretical because every source has some spectral spread. The simplest lasers to manufacture are multilongitudinal-mode lasers, which produce light at a number of different wavelengths spread over 5 to 10 nm or more. More sophisticated designs (distributed feedback, external cavity, and so on) reduce the spectrum to a single line, which is narrow enough so that chromatic dispersion can be ignored in most applications; these devices, however, are expensive.
2. A more common way is to choose a transmitter center wavelength and fiber design so that the effect is minimized. Silica-based optical fiber has low loss at 1,300 nm and even lower loss at 1,550 nm, making these good operating wavelength regions. Close to 1,300 nm, two signals at 10 nm apart in wavelengths (a typical laser line width) will have little difference in velocity. In contrast, a laser pulse with a 1-nm

Figure 2-13 Temperature-dependent chromatic dispersion.



spectral width centered at 1,550 nm will be broadened by 16 picoseconds after traveling 1 km. If the spectral width is 10 nm and the fiber length is 10 km, the pulse will be broadened by 1,600 picoseconds (1.6 nanoseconds). Negative values of chromatic dispersion have the same effect as positive values (the positive or negative sign indicates if longer or shorter wavelengths travel faster).

3. When utilizing WDMs, multiple wavelengths travel down a fiber by design. Here the approach to address dispersion is to utilize a certain amount of dispersion compensation fiber (DCF) at the receiving end. By properly selecting a terminating fiber that has the opposite effect of the transmission fiber in terms of dispersion, you can reduce the impact of the dispersion phenomenon. The add-on fiber takes the rays that traveled faster and slows them down more than the rays that traveled slower. Unfortunately, a rather long spool of fiber is needed, which is placed in a “pizza box.” Such spools could have several kilometers of fiber. Pizza boxes have to be selected based on the length of the transmission link. This requirement constitutes a problem because add-on fiber further impacts attenuation—the added fiber must be taken into account in the power budget. A relatively new approach is to use tunable dispersion compensators (see Figure 2-14 for examples).[5]

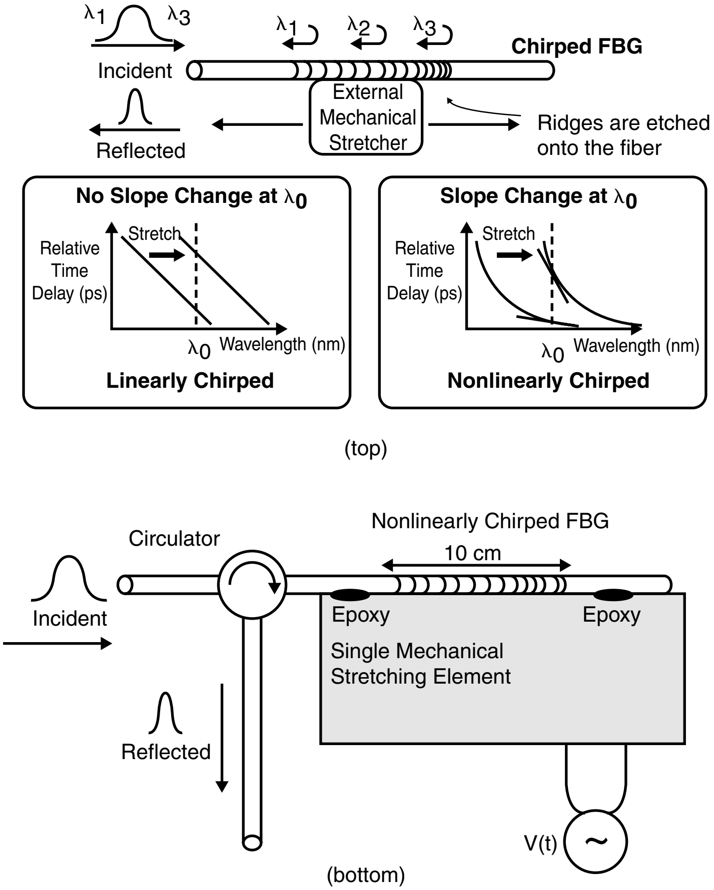


Figure 2-14 Dispersion compensation without DCF.

Top: Dispersion Compensation Using a Chirped Fiber Bragg Grating
Bottom: Tunable Dispersion Compensator

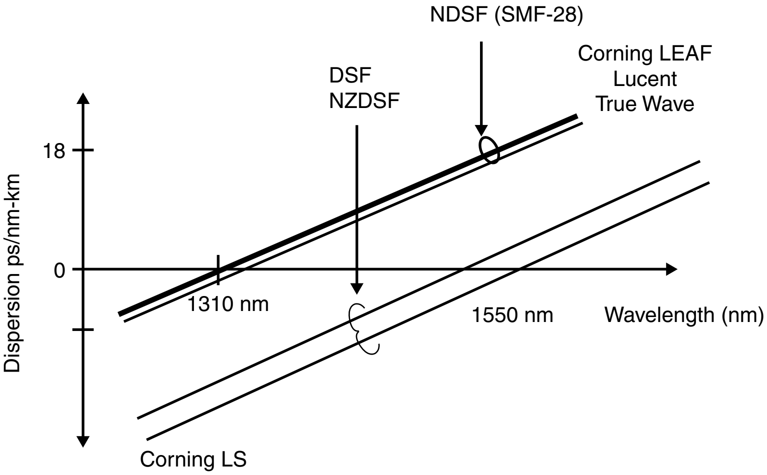
The bottom line is that chromatic dispersion can be managed, even at higher speeds, but only at the cost of system and power.

Returning to chromatic dispersion on non-WDM-used fiber, the optimal center wavelength for the transmitter is the point where chromatic dispersion is zero. This wavelength is called λ_0 or the zero-dispersion point (ZDP). This value varies from optical fiber to optical fiber and is an important parameter for fiber manufacturers to control. Another parameter of interest is the slope of chromatic dispersion near λ_0 called the zero-dispersion slope (S_0). S_0 is a measure of how rapidly chromatic dispersion increases as you move away from ZDP.

ZDP occurs naturally in pure silica glass at 1.27 microns (see Figure 2-15). Because, in principle, single-mode fibers work with a single coherent wavelength, one way to exploit ZDP is to find a laser that emits light at 1.27 microns. However, the search for the greatest overall efficiency entails other factors. Usually, glass waveguides suffer from losses due to Rayleigh scattering, which occurs because of density and compositional variations within the glass. (Rayleigh scattering lessens as wavelengths grow longer.) To take advantage of this, dopant materials can be added to glass until its ZDP is shifted into the range between 1.3 and 1.6 microns. Many formulations of glass reach their lowest possible absorption in this range (below this range, molecular thermal scattering becomes a problem). Manufacturing must be precise to avoid impurities in glass that increase absorption.

Ideally, in non-WDM situations, the laser would operate at this wavelength so the bandwidth would be very high; however, due to manufacturing variations and aging, laser-operating wavelengths cover a broad range of operating wavelengths. For a given transmission path length, a single-mode fiber-optic system exhibits less pulse broadening than multimode systems, provided that the central wavelength is sufficiently close to the fiber's ZDP.

Figure 2-15
Dispersion behavior of two fiber types.



For long-distance applications, it is desirable to operate at 1,550 nm because of lower attenuation compared to 1,300 nm. However, a simple step-index single-mode fiber shows a large chromatic dispersion at 1,550 nm. It is possible to design more sophisticated index profiles so that the point where chromatic dispersion is zero falls near 1,550 nm instead of 1,300 nm. This optical fiber is called a dispersion-shifted fiber (DSF). This fiber may give significantly longer unrepeated distances at higher information transmission rates than the unshifted fiber when conventional LDs or LEDs are used. Segmented-core fiber is a fiber that is designed to operate as a low-loss single-mode fiber at 1,300 and 1,550 nm while maintaining a zero-dispersion wavelength in the 1,550-nm window. In addition, this fiber performs as a high-bandwidth few-mode fiber near 800 nm where inexpensive light sources are available. Segmented-core fiber enables mode equalization over a large range of wavelengths. This fiber can be used in high-performance long-haul or in subscriber-loop applications at longer wavelengths. Figure 2-16 contains taxonomy of the key fiber types currently available.

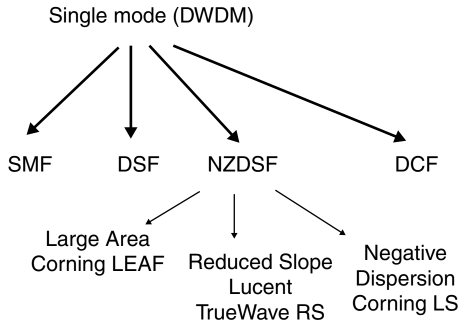
The six major generations of fiber-optic systems spanning the past 20 years are as follows:

- **First** Multimode (step index)—850 nm
- **Second** Multimode (graded index)—1,300 nm
- **Third** Single mode (pre-SONET/SDH)—1,300 nm
- **Fourth** Single mode (dispersion shifted) (SONET)—1,550 nm
- **Fifth** WDM systems; also OC-192 (SONET)—1,550 nm
- **Sixth (latest)** Dense WDM (DWDM) systems; also OC-768 (SONET) —1,550 nm

A number of fiber types and the behaviors of rays as they travel down the fiber are identified in Figure 2-17. Research is continuing on nonsilica glass, which theoretically offers much less loss. Plastic fibers are also available for intrabuilding use in applications such as LANs. The attenuation of this fiber, however, is very high; hence, the majority of commercially available Gigabit Ethernet (GbE) and 10-GbE LANs use multimode silica fiber rather than plastic.

In conclusion, chromatic dispersion can be managed. At higher speeds, the issue becomes more pressing (particularly at 40 Gbps on NDSF; operation at OC-768 is 16 times more sensitive to the problem than operation at OC-192). Managing chromatic dispersion imposes a cost and a loss premium.

Figure 2-16
Fiber taxonomy.



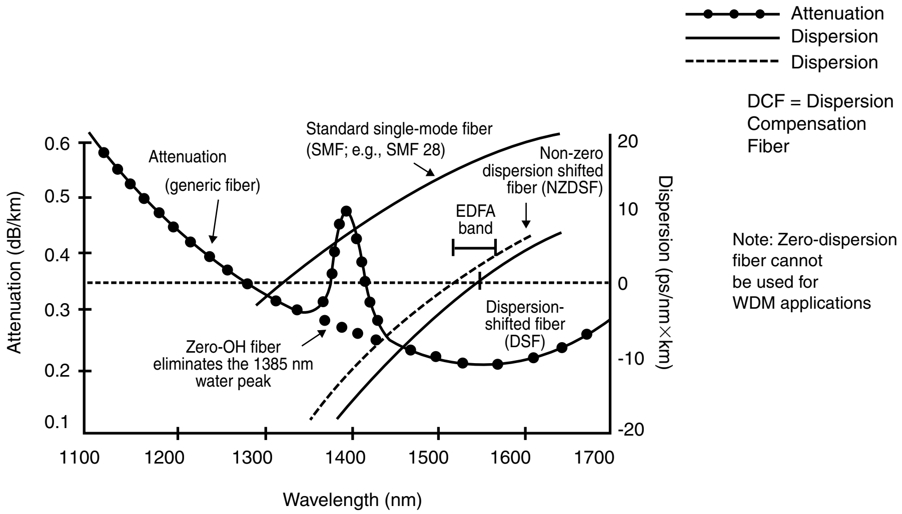
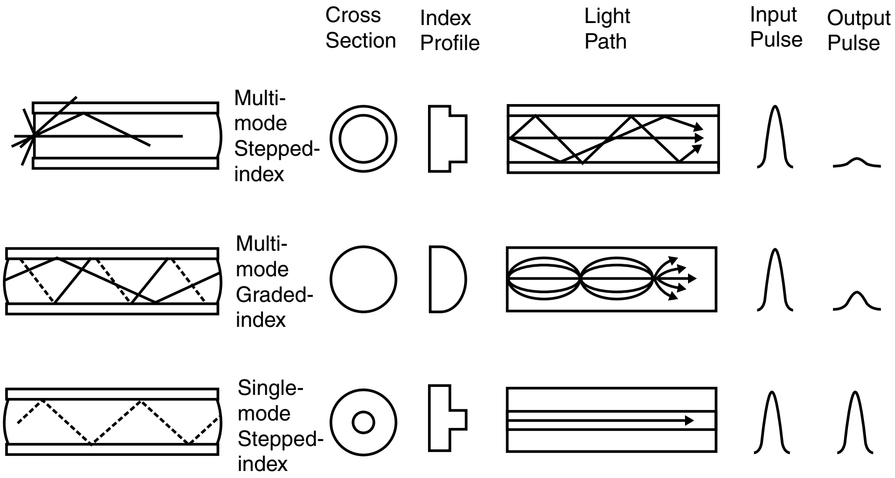


Figure 2-17
Behavior of the rays as they travel down the fiber.



2.1.5 Single-Mode Fibers

As noted, single-mode fibers are designed so that only one mode propagates. Here the information transmission capacity is limited by the phenomenon of chromatic dispersion, which, as discussed previously, is a result of the wavelength-dependant velocities of propagation in the fiber material. Conventional single-mode fibers (known by the siglum SMF-28) have their minimum chromatic dispersion (that is, their maximum transmission capacity) near 1.31 μm.

Conventional single-mode fibers are classified as either matched-cladding or depressed-cladding designs. In the former type, the cladding has the same index of refraction from the core-cladding interface to the outer surface of the fiber. In the latter type, there are two layers of cladding material. The index of refraction of the inner cladding, which is adjacent to the fiber core, is lower than that of the outer cladding (typically undoped silica).

Single-mode fibers are characterized by the mode field diameter (MFD) and by the cutoff wavelength. The MFD is a measure of the width of the guided optical power's intensity distribution in the core and cladding. Typical values for the MFD are in the 10.0-μm range. The cutoff wavelength is the wavelength below which a single-mode fiber propagates more than one mode and above which only a single (fundamental) mode can operate.

DSF is designed so that the wavelength of minimum dispersion is shifted to 1.55 μm. At this wavelength, the attenuation can be significantly lower than that at 1.31 μm. Single-mode fibers can also be designated as dispersion-flattened fibers. Dispersion-flattened single-mode fibers are attractive. The dispersion is low in both the 1.31-μm window and 1.55-μm window.

2.1.6 Multimode Fibers

Multimode fibers support dozens of propagating modes, each of which travels at a slightly different velocity. These fibers tend to be used in GbE and 10-GbE LAN applications, although they also can utilize single-mode fiber. The core diameter of the multimode fiber normally used is 62.5 μm, but other special designs of multimode fibers exist. Multimode fiber can be step-index or graded-index fiber, although most of the fiber deployed is step-index fiber. As we've seen, graded-index multimode fiber reduces pulse distortion by nearly equalizing the velocities of modes propagating in the fiber. With graded-index fiber, the index of refraction decreases almost parabolically from the center of the core to the boundary between the core and cladding material. However, the equalization of mode velocities is not perfect, and the information transmission capacity of multimode fibers is substantially less than that of single-mode fibers.

Roughly speaking, 850- and 1,310-nm multimode systems can operate at distances from a few hundred meters to a couple of kilometers; see Table 2-2[6] for examples. A 1,310-nm Fabry-Perot (FP) laser single-mode system can range from 10 to 40 km; Distributed Feedback (DFB) lasers can reach up to 100 km with good fiber. If the distance spanned exceeds these limits, the light must be amplified directly (with an erbium-doped fiber amplifier [EDFA]) or converted to electricity, regenerated, amplified, and converted back to light.

2.1.7 Optical Fiber Specs

Fiber cable has traditionally been codified using a three-code field as follows:

$M_1M_2\ S_1S_2S_3S_4S_5S_6$ NNN

M_1M_2 is a manufacturer code; NNN is a code indicating the number of working fibers in the cable. The meaning of the S field is discussed next.

S_1 defines the mode of the fiber and wavelength at which the attenuation is specified, as shown in Table 2-3.

Character S_2 defines the transmission performance of the cable. S_2, in conjunction with S_1,[7] indicates the maximum attenuation or the minimum bandwidth (in MHz×km) of several different fibers, as indicated in Table 2-4. The objective for the average attenuation coefficient of all fibers in a cable should be no greater than 80 percent of the maximum coefficient given in the table.

Table 2-2

Distances Supported by Multimode Fiber

Application	**Wavelength**	**Distance**
10-Mbps Ethernet, 10Base-FL	850-nm MM	2,000 m
100-Mbps Fast Ethernet, 100Base-FX	1,310-nm MM	2,000 m
100-Mbps Fast Ethernet, 100Base-SX	850-nm MM	300 m
GbE, 1000Base-SX	850-nm MM	220 m/550 m*
GbE, 1000Base-LX	1,310-nm SM	550 m

*62.5μm/50μm fiber

Table 2-3

Fiber Mode and Attenuation Wavelength Defined by S_1

	Usable Wavelengths (µm)				
Fiber Type	**0.85**	**1.30**	**1.55**	**0.85 and 1.30**	**1.30 and 1.55**
Single mode	—	1	—	—	2 or 3
Dispersion shifted	—	—	6	—	7
50/125 multimode	K	L	—	M	—
62.5/125 multimode	P	Q	—	R	—
85/125 multimode	T	W	—	Y	—

Power

$P = iV = i^2R = V^2/R$

Decibel

This standard unit is used to express transmission gain or loss, and relative power level. One decibel (dB) is 1/10 of a bel. It is used to describe the ratio of the power at any point in the transmission line to a reference of 1 milliwatt of input. (The ratio expresses decibels above or below this reference level of 1 milliwatt.)

DBM is used when a power of 1 milliwatt is the reference level.

dB indicates the ratio of power output to power input:

$dB = 10 \log_{10} (P_1/P_2)$

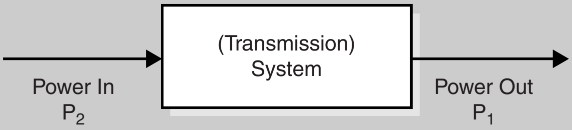


Table 2-4

Character S_2

Character S_2 for Conventional Single-Mode Fiber

Character	1.30 μm Attenuation (dB/km)
4	0.40
5	0.50
6	0.60
7	0.70

Character S_2 for DSF

Character	1.30 μm Attenuation (dB/km)
2	0.40
3	0.50

Character S_2 for DSF

Character	1.55 μm Attenuation (dB/km)
2	0.25
3	0.30

Character S_2 for Multimode Fiber, 50/125

Minimum		Maximum Attenuation (dB/km)				
Bandwidth	(MHz×km)	0.85	μm:	2.6	3.0	3.5
0.85 μm	1.30 μm	1.30	μm:	0.8	1.0	1.6
300	400			G	N	U
500	800			F	M	T
600	1,000			E	L	S
NA	1,200			D	K	R
NA	1,500			C	J	Q

Character S_2 for 62.5/125 and 85/125 Multimode Fiber

Minimum		Maximum Attenuation (dB/km)				
Bandwidth	(MHz×km)	0.85	μm:	2.6	3.0	3.5
0.85 μm	1.30 μm	1.30	μm:	0.8	1.0	1.6
150	300			F	M	T
200	600			E	L	S
250	800			D	K	R
NA	1,000			C	J	Q

For single-mode fibers ($S_1 = 1, 2$, or 3), the S_2 character defines the specified maximum individual-fiber attenuation in the cable. The maximum individual-fiber attenuation at 1.55 μm depends on the first character, S_1. When $S_1 = 1$, then the 1.55-μm attenuation is not specified; when $S_1 = 2$, then the maximum individual-fiber attenuation at 1.55 μm is the same as at 1.30 μm, as shown in Table 2-4; when $S_1 = 3$, then the maximum individual-fiber attenuation at 1.55 μm is 0.10 dB/km less than at 1.30 μm.

For dispersion-shifted single-mode fibers, which are optimized for use at 1.55 μm, $S_1 = 6$ or 7. When $S_1 = 6$, only the attenuation at 1.30 μm is specified; when $S_1 = 7$, the maximum individual-fiber attenuations at 1.30 and 1.55 μm are defined, as shown in Table 2-4.

For 50/125 multimode fibers, with S_1 = K, L, or M, the S_2 character denotes both the minimum allowable bandwidth (in MHz×km) and the maximum attenuation (in dB/km). Because of the large number of possible combinations of attenuations and bandwidths at the two wavelengths, only a few of the more common combinations are coded, as shown in Table 2-4. Other combinations not shown are coded as S_2 = X. For S_1 = K, only the 0.85-μm attenuation and bandwidth values apply (the upper of the two attenuation headings and the left bandwidth column). For S_1 = M, both 0.85- and 1.30-μm values apply. The code also covers 62.5/125 and 85/125 multimode fibers. As with 50/125 fibers, only a few of the most common attenuation and bandwidth combinations are coded. Combinations not shown in the table are coded S_2 = X. For S_1 = P or T, only the 0.85-μm values apply. For S_1 = Q or W, the 1.30-μm values apply. For S_1 = R or Y, both values apply.

2.2 Fiber Splicing

Fiber rolls may be of the order of 2 to 4 km; hence, longer metropolitan or long-haul links will entail splices. Splices are also involved in the building spur to metro access ring. A splice is a semipermanent or permanent connection of two fibers. Semipermanent splices are made with the aid of a mechanical device holding the fiber ends together. Mechanical splices can be reentered, if necessary, without breaking or cutting the fiber. These splices can also be made without the use of expensive splicing tools. The fibers are simply held together with a connector that aligns them. The alignment is usually accomplished by adjusting the outer surface of the fiber. Mechanical splices are normally used to reduce splicing times. They involve some risk of temperature-induced misalignments that introduce variations in loss. Connector losses for a 12-fiber array splice have a loss of about 0.4 dB.

Permanent splices are made either by electric-arc fusion or by chemical bonding. Electric-arc fusion splicing is the most widely used splicing method. The idea behind this method is to heat the fibers and fuse them together. Chemically bonding fibers together is the procedure of gluing two fiber ends together chemically. The adhesive material is typically a synthetic polyester. Because the fiber-end faces do not come in direct contact in this case, the adhesive must serve as an index-matching material to provide a consistent refractive index at the joint. Fusion or bonding undertaken under controlled conditions produces low-loss splices. Losses of 0.05 dB can be achieved with these methods.

2.3 Additional Details on Limitations of Fiber Links

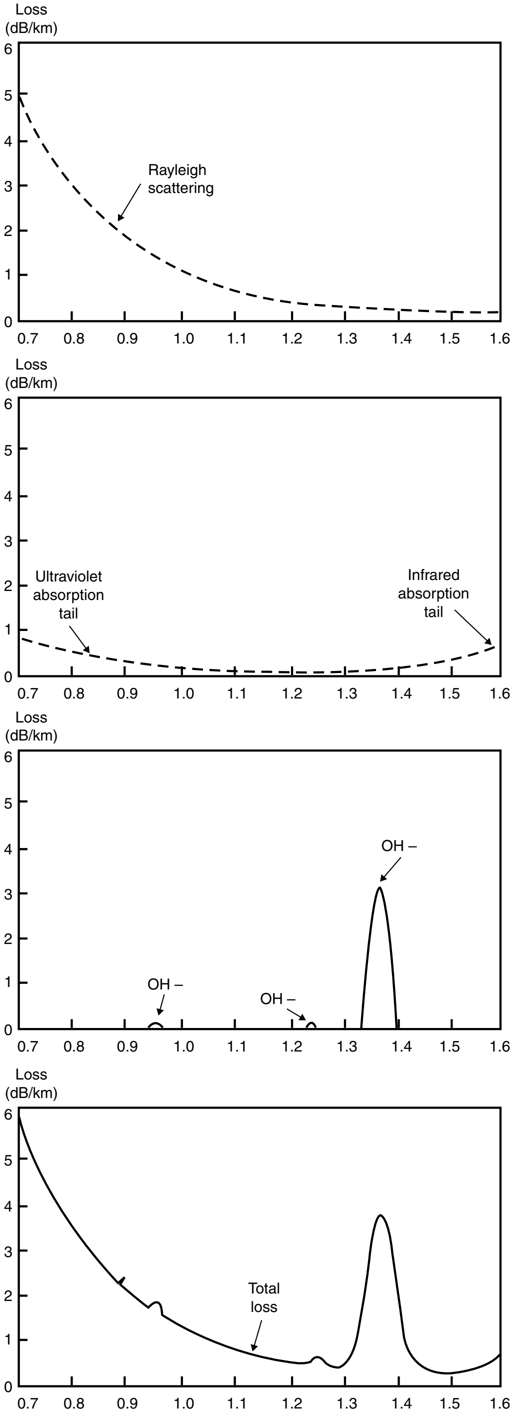
Fiber-optic transmission systems contain fundamental limitations on both the distance and the bit rate at which information can be transmitted. As discussed in the previous section, these limitations are determined by the attenuation of the signal and by the spreading of the digital pulses caused by modal and chromatic dispersion within the fiber media. For single-mode systems running up to 0.5 Gbps, link length before regeneration is limited by loss rather than by dispersion. At higher bit rates, however, the length is limited by dispersion. Length issues are generally more of a concern for long-haul systems; however, there are situations where they also impact metropolitan applications.[8]

2.3.1 Fiber Attenuation

Attenuation of a signal in silica-based fiber at long wavelengths is due to three major causes: Rayleigh scattering, absorption, and bending loss. Figure 2-18 shows the scattering/absorption issue.

- Rayleigh scattering is caused by the microscopic nonuniformity of glass and its refractive index. A ray of light is partially scattered into many directions, and therefore some of the light energy is lost. Because the structure of glass is much finer than the wavelength, the attenuation due to scattering decreases with wavelength. Attenuation is proportional to $1/\lambda^4$.

Figure 2-18
Attenuation factors.



- Absorption is caused by interactions of the signal with impurities in the silica glass, causing unwanted radiation of the signal. Loss peaks at a wavelength of about 1.39 μm, caused by OH-radical impurity (see Figure 2-10). OH-radical impurity must be controlled to prevent loss spreading into the usable wavelength bands. Losses for high-quality fiber are typically less than 0.40 dB/km and 0.22 dB/km at 1.31 μm and 1.55 μm, respectively. New fibers have been developed that eliminate the 1,385-nm water peak.
- Attenuation is also caused by bending. There are two forms of bending: macrobending and microbending.[9] Macrobending is loss that is attributable to macroscopic deviations of the axis from a straight line. Microbending loss occurs from the sharp curvatures involving local axis displacements of a few micrometers and spatial wavelengths of a few millimeters. These bends result from optical fiber coating, cabling, packaging, installation, and aging.[10]

2.3.2 Dispersion and Zero-Dispersion Point (ZDP)

Intramode dispersion can be subdivided into material dispersion and waveguide dispersion. Material dispersion refers to the wavelength dependence of the fiber's refractive index and the associated differences in the speed of light. Telecom system light sources have a spectral width of at least 1 nm (see Figure 2-19). As we've noted previously, different components within this spectrum travel at slightly different speeds. The slope of material dispersion is positive.

Waveguide dispersion causes pulse broadening in a fiber channel due to spectral components traveling at slightly different speeds, a wavelength dependence related to the geometry of the fiber, and the shape of the index profile. The geometry of the fiber causes the propagation constant of each mode to change if the wavelength changes. The slope for waveguide dispersion can be negative. This can be used to shift the point of zero dispersion or to flatten the dispersion.

Single-mode silica fibers have one of two refractive-index profiles: matched-cladding fiber and depressed-cladding fiber. In depressed-cladding fiber, the index of the cladding next to the core is less than that of pure silica; this is done to better confine the optical power distribution to the core (refer once more to Figure 2-5). The ZDP for these fibers occurs at about 1.3 μm; the ZDP can be shifted by changing the index profile to a triangular

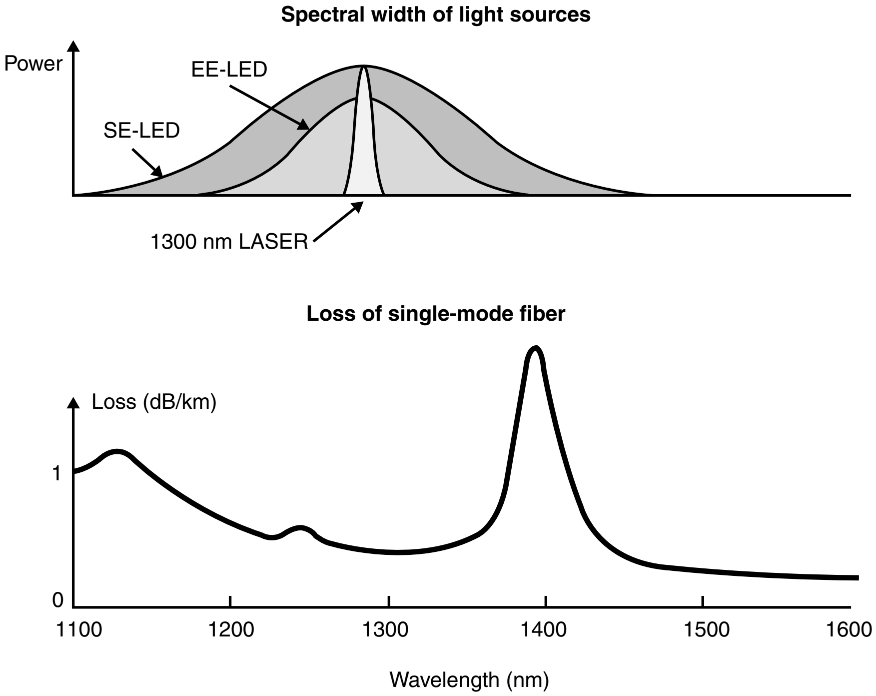


Figure 2-19 Spectral scope of sources.

shape, making the zero occur near 1.55 μm. As Figure 2-18 suggests, this is where the lowest loss can be obtained.

By trading off waveguide dispersion against material dispersion over a range of wavelengths, a dispersion-flattened fiber can be developed. Flattening the dispersion enhances the bandwidth properties of the fiber over a greater range of wavelengths. (The index profile for this design is shown in Figure 2-5[f].) Figure 2-20 plots the dispersion for three kinds of fiber: conventional, flattened, and shifted. Figure 2-21 expands upon the trade-off by showing the dispersion values of conventional fiber compared with the fiber's loss. As you can see, conventional fiber loss is minimal at about 1.55 μm and has zero dispersion near 1.3 μm.

Figure 2-22 shows the bandwidth-length product to be greater at the ZDP; however, the bandwidth depends on the spectral width of the source (such as an LD or LED). For example, with a source spectrum $\delta\lambda$ of 40 nm (such as for an LED transmitter), the bandwidth-length product for the fiber at the ZDP is about 10 GHz/km, and for a source spectrum of 2 nm

Figure 2-20
Single-mode fiber dispersion as a function of wavelength.

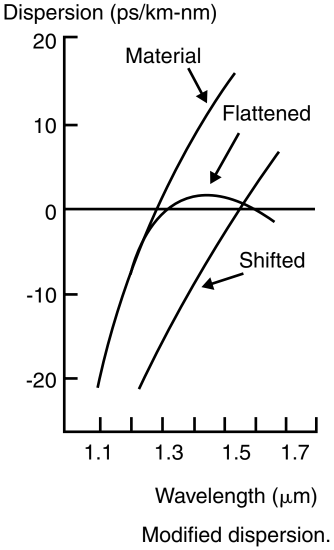
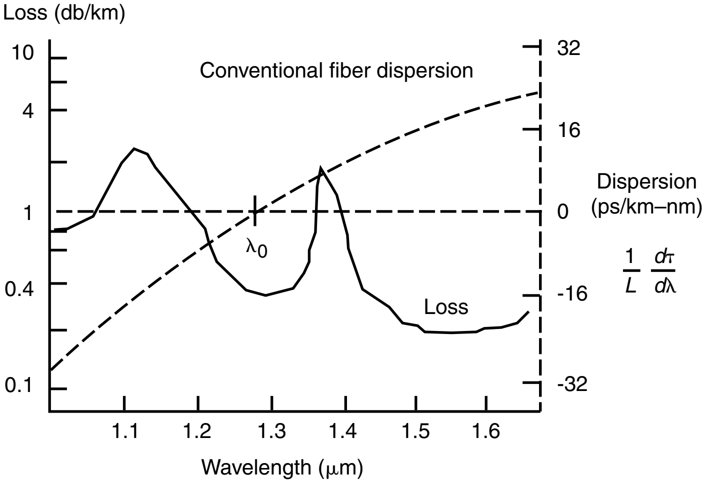
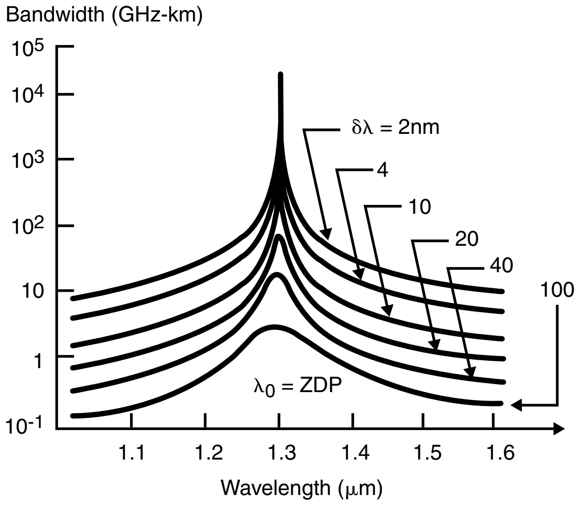


Figure 2-21
Characteristics of conventional fiber.



λ_0 = ZDP
$\delta\lambda$ = Line width of source spectrum

Figure 2-22 Single-mode fiber transmission characteristics.



λ_0 = ZDP
δλ = Line width of source spectrum

(such as for an LD transmitter), the product is over 100 GHz/km. For an equivalent 3-dB power penalty at a BER of 10^{-9}, the bandwidth-length product in terms of GHz×km is determined by the following:[10]

$$(B)\,(L)\begin{cases} \leq \dfrac{340.5}{(\sigma)\left|\dfrac{d\tau}{d\lambda}\right|} \text{ away from the ZDP} \\[2ex] \leq \dfrac{11207}{(\sigma)^2\left|\dfrac{d^2\tau}{d^2\lambda}\right|} \text{ at the ZDP} \end{cases}$$

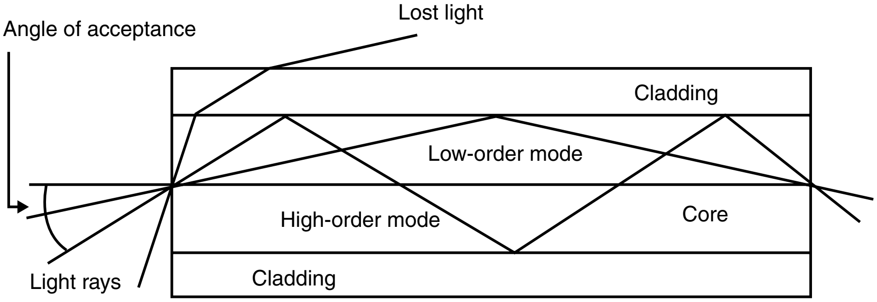
(B) = Bandwidth (Gbps)
(L) = Length (km)
σ = Root mean square half-width of the source spectrum (nm)
τ = Refractive index

Multimode fibers cause source-pulse broadening that is due to the different propagation speeds of the different modes. Because this usually has a larger effect than chromatic dispersion, multimode systems exhibit more pulse broadening. The different propagating modes of light carrying the information effectively zigzag down the fiber core, angling from side to side because of the total internal reflection between the core and its surrounding cladding (see Figure 2-23). The zigzag path taken by light in the higher traverse modes means that those rays follow longer paths than the rays of the fundamental transverse mode that travel directly down the center of the fiber's core. Because of these different path lengths, pulses following the transverse modes are gradually spread. Those along the higher modes arrive at the receiver somewhat later than the rays of the fundamental mode. Thus, even though in each digital pulse all the rays have left the source at the same time, modal dispersion broadens the pulse as it travels down the fiber to the detector. The amount of broadening is a function of length and increases as distance increases. This phenomenon increases the BER at the receiving end.

2.3.3 Polarization Mode Dispersion (PMD)

PMD is caused by slight fiber asymmetry. During the perform fabrication or during the drawing process, a nonperfectly circular core and/or cladding may result (see Figure 2-24). PMD is the result of the accumulation of weak birefringence along the various fiber spans, enabling a link of any length exceeding 2 or 4 km to be comprised of spliced fibers that originate from different manufacturing events. An input light pulse of an initial polarization will decompose into two pulses with two different polarizations separated

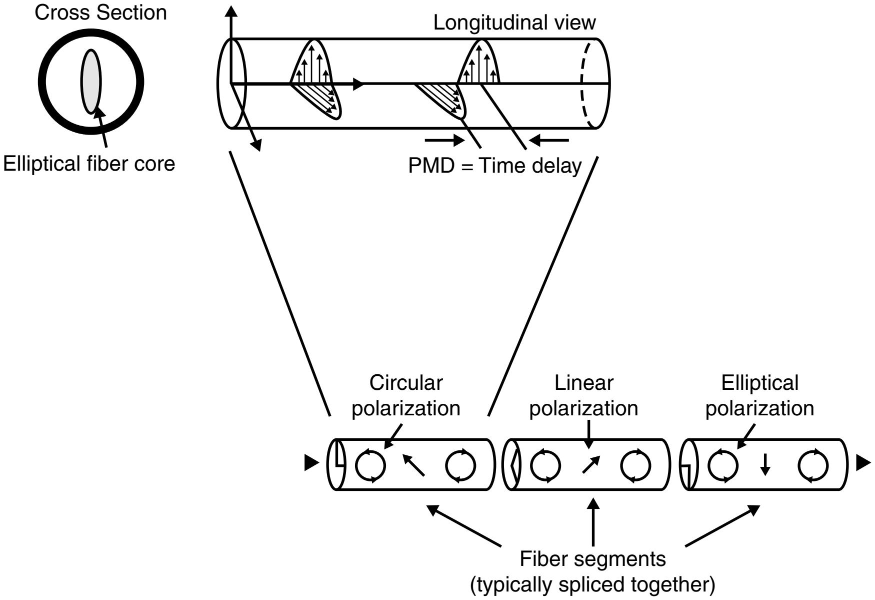
Figure 2-23
Zigzag rays in a multimode fiber.



(Three modes are shown in the figure; in actuality hundreds of discrete modes typically propagate in a single fiber.)

Low-order modes, those launched into the fiber at small angles with respect to the axis, travel a shorter ray distance to reach a given point along the fiber than do the high-order modes, those launched at large angles.

Figure 2-24
PMD.



in time. The time separation leads to decreased system margin or even outages. This issue becomes more pronounced and problematic at higher speeds (such as OC-192 and OC-768).

In addition to noncircular cores and cladding from the manufacturing process, bends and/or twist-induced stresses also introduce PMD. Furthermore, PMD also depends on the temperature of the fiber, sometimes making transmission problematic on a seasonal basis. System components such as isolators and circulators also have a bearing on the issue. The problem is more pronounced on older fibers and systems. At this time, fiber can be drawn in a manner that minimizes PMD; also, components can be manufactured that keep PMD low. Older fibers have PMD in the range of 0.5 ps/km$^{1/2}$; newer fibers range around 0.1 ps/km$^{1/2}$. The expectation is that this figure will decrease to 0.05 for future fibers. Forward error correction (FEC) methods may also be used as an effective tool against nonlinear effects. Tunable PMD compensators are also being developed at this time.

2.3.4 Other Impairments

Noise is measured in terms of the optical signal-to-noise ratio (OSNR). Each (optical) amplifier along a fiber path adds noise, giving rise to the need for regeneration every so many kilometers. Longer spacing between amplifiers leads to degraded OSNR figures. In DWDM systems, the larger the number of channels, the higher the amplification requirement to maintain an acceptable OSNR. Higher speed also requires improved OSNR (for example, 10-Gbps systems require 6 dB more OSNR than 2.5-Gbps systems).

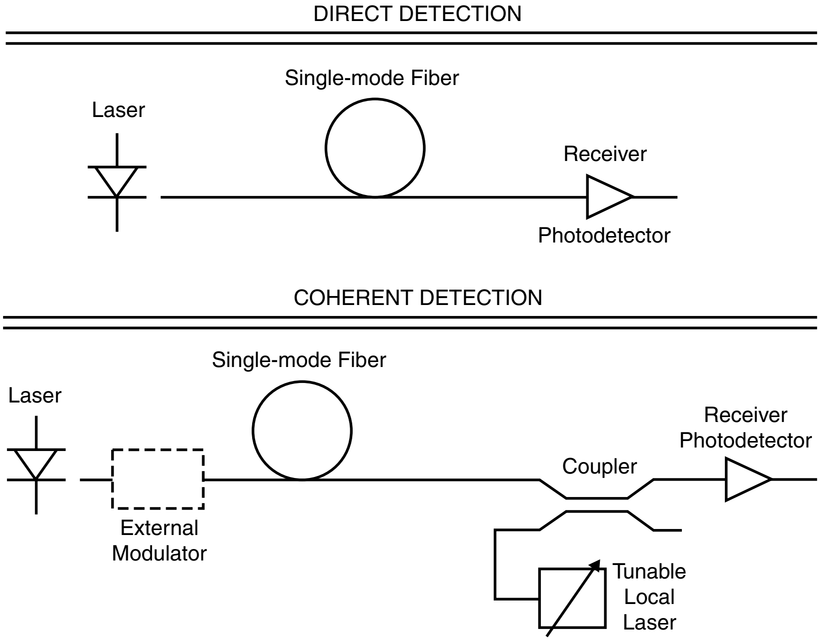
Nonlinearities give rise to the following other problems that have to be dealt with appropriately:

- Stimulated Raman Scattering (SRS)
- Stimulated Brillouin Scattering (SBS)
- Four-wave mixing
- Self-phase modulation
- Cross-phase modulation

2.4 Transmission Approaches

The two basic approaches to optical transmission—noncoherent and coherent—are described in Figure 2-25.

Figure 2-25 Transmission methods.



2.4.1 Noncoherent Transmission

Noncoherent methods involve on-off keying and the direct detection of a signal; this is a variant of amplitude modulation (AM), which is specifically a unipolar digital scheme of amplitude shift keying (ASK). The amount of information transmitted is based on the source/reception clocking. Building on this basic method, several noncoherent optical channels can be operated over the low-loss/low-dispersion ranges of an optical fiber. This requires filters to separate the optical channels. Such devices (filters) are called WDMs. Initially, such multiplexers were designed to support two channels in the 1.30-μm wavelength region and three channels in the 1.50-μm region. Commercial products now support several dozen channels and laboratory tests have demonstrated the possibility of using several hundred channels of 1 to 10 Gbps, utilizing narrow spectrum lasers and direct detection.

Even this multiplexing, however, does not use the available fiber bandwidth to its maximum potential (100 channels at 10 Gbps would equate to 1 terabit per second, or 0.001 petabits per second). First, there are several limitations such as source spectra, multiplexer filter technique, modulation-induced effects such as frequency chirp and mode-partition noise, and fiber dispersion characteristics.[10] Secondly, the available fiber bandwidth can be estimated from looking at wavelength regions having low loss (less than 0.5 dB/km or so), such as from 1.27 to 1.35 μm and from 1.48 to 1.60 μm. These regions provide a total wavelength band of 200 nm and correspond to a frequency bandwidth of about 29 terahertz yet it is used to support the few (dozen) noncoherent channels. Coherent transmission can provide many more channels. Assuming a 10-bits-per-hertz modulation scheme, the 29 terahertz could support 0.3 petabits per second (0.3×10^{15} bps).

2.4.2 Coherent Transmission

The available fiber bandwidth could be used more efficiently if coherent transmission (full-frequency modulation) is used. The optical frequencies could be down-converted to lower radio frequencies and processed electronically with existing technology for improved sensitivity, filtering, and equalization. This, however, requires several stringent design considerations: a laser with a very narrow line width or single-frequency spectrum and single longitudinal mode, very high frequency and phase stability, automatic polarization correction, automatic frequency control of the source and of the local receiver laser, and so on. Because of these stringent design requirements, coherent transmission methods have not yet seen commercial deployment.

2.5 Design Parameters

This section summarizes some design parameters for typical telecommunication applications.[10]

2.5.1 Attenuation Coefficient Versus Wavelength

For single-mode fibers, the coefficient between 1.285 and 1.330 μm should not exceed that at 1.31 μm by more than 0.1 dB/km. This is to provide a low-loss transmission window.

2.5.2 Water Peak Attenuation

The coefficient at 1.38 μm should not exceed an objective of 3 dB/km to control this contamination.

2.5.3 Attenuation Uniformity

The attenuation of a continuous fiber is expected to be distributed uniformly throughout its length so that there is no discontinuity in excess of 0.1 dB. Splices made in manufacturing are typically at least 1 km apart; these splices must meet the attenuation requirements as part of the fiber. This implies that bonded or fused splicing with accurate concentricity is necessary to produce splice losses on the order of 0.05 dB or less.

2.5.4 Attenuation with Bending

For multimode fibers, the attenuation introduced when 100 turns are wound on a 75-mm diameter mandrel (including the substance's own attenuation of 23.6 meters of fiber) is expected not to exceed 0.5 dB at any usable wavelength. For single-mode fibers, the attenuation of these 100 turns is expected not to exceed 1 dB at 1.55 μm. Depressed-cladding fiber is more resistant to these losses due to a smaller MFD (better optical-power containment). Because these losses increase as wavelength increases, the depressed-cladding design is better suited for operations at longer wavelengths such as 1.55 μm.

2.5.5 Chromatic Dispersion (Single-Mode Fiber)

For conventional single-mode fibers, the zero-dispersion wavelength, λ_0, typically is between 1.298 and 1.322 μm, with a nominal value of 1.31 μm. The maximum value of the dispersion slope at λ_0, S_{omax}, is typically less than 0.095 ps/(km×nm^2).

For dispersion-shifted single-mode fibers, the zero-dispersion wavelength is between 1.54 and 1.56 μm, with a nominal value of 1.55 μm. In addition, S_{omax} is typically no greater than 0.07 ps/(km×nm^2).

2.5.6 Multimode Bandwidth

The minimum bandwidth is to be specified at the wavelength of intended use either by the end-to-end bandwidth requirement of the cable span or by an individual-reel bandwidth requirement. Typical minimum bandwidth categories for the three multimode fiber designs were given in Table 2-4.

2.5.7 Cutoff Wavelength (Single-Mode Fiber)

The cutoff wavelength of a cabled fiber, λ_{cc}, should be between 1.25 and 1.03 μm. As an alternative, manufacturers may specify the cutoff wavelength of an uncabled fiber, λ_{cf}, which lies between 1.29 and 1.13 μm. You should operate above the cutoff wavelength to prevent multimode behavior in a single-mode fiber, causing modal noise and pulse broadening; hence, you want to have λ_{cc} as low as possible.

2.5.8 Mode Field Diameter (MFD) (Single-Mode Fiber)

The nominal MFD for conventional single-mode fibers is between 8.7 and 10 μm. For dispersion-shifted single-mode fibers, the nominal should be between 7 and 8.7 μm. However, these values depend on the measurement technique used.

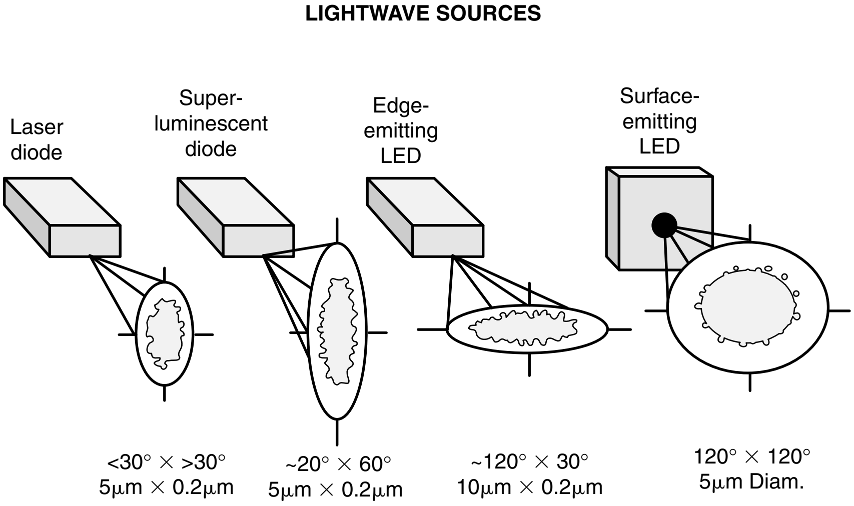
2.6 Fiber-Optic Transmitters

Optical transmitters convert digital electrical signals into pulsed optical signals; furthermore, they couple these signals into the fiber waveguide for transmission. The two major kinds of semiconductor transmitters in use are LEDs and LDs. These transmitters have the necessary design characteristics of small size, conversion efficiency, coupling efficiency, speed, electrical amenability, environmental resistance, and reliability. LDs are typically used with single-mode fibers, and LEDs are typically used with multimode fibers (however, other combinations are possible). Figure 2-26 shows the photonic footprint of these and other sources. Efforts have been underway for a number of years to integrate the optics and electronics monolithically in an electro-optical chip in order to provide for higher-bit-rate transmission.

2.6.1 Light-Emitting Diodes (LEDs)

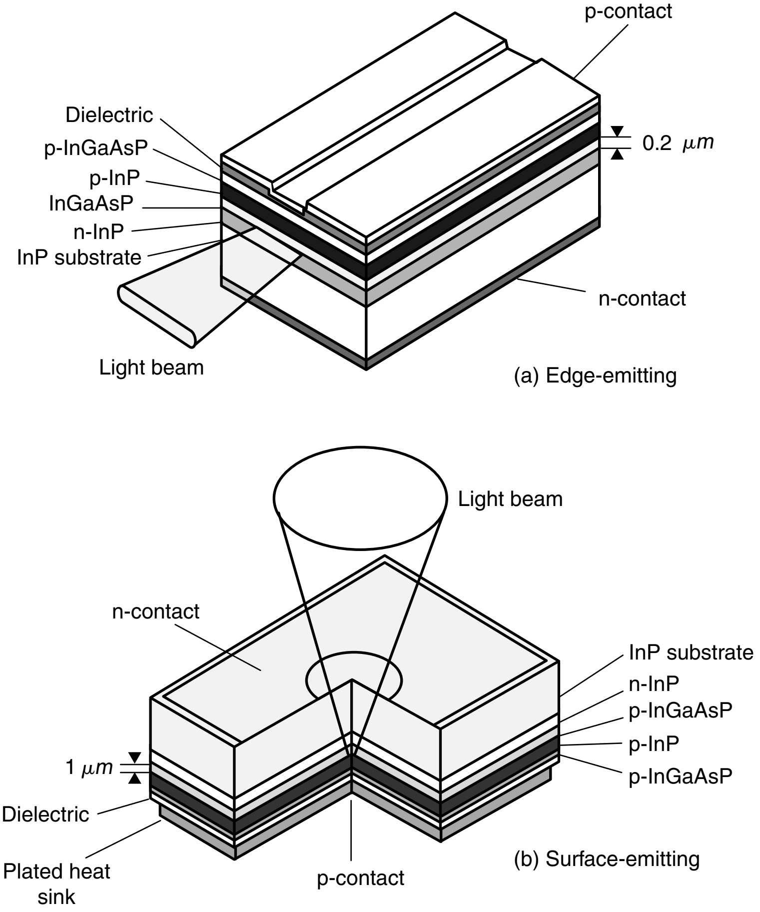
LEDs come in two types: surface emitters and edge emitters. The surface-type emission pattern is typically 120°×120°. Light emission occurs at currents of 100 mA. The peak power coupled into multimode fiber is about −10 to −20 dBm; the power coupled into single-mode fiber is about −27 to −37

Figure 2-26
Light sources.



dBm. The spectral (noise) width of the emission is about 35 to 50 nm for short wavelengths and about 80 to 120 nm for long wavelengths. The edge-emitting diode has a more focused emission pattern, typically 120°×30°, and so the power coupled into the fiber is up to 10 dB greater than that for surface emitters. Over the band of interest, spectral width is about 30 to 90 nm. The spectral width of the LED source and the dispersion characteristic of the fiber limit the bandwidth that can be used. LED structure is similar to lasers and care must be taken to prevent unwanted laser action. Figure 2-27 depicts edge- and surface-emitting LEDs. The LEDs are operated with a small forward bias, that is, low light emission, to overcome a turn-on delay. Also, because the output optical power varies with DC voltage and

Figure 2-27
LED structures.



temperature, compensation of the LED is required. A thermoelectric cooler may also be used to stabilize the LED temperature over a range of temperature differences between the LED and the ambient.

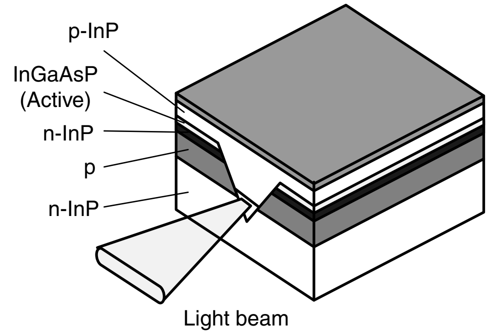
2.6.2 Laser Diodes (LDs)

Semiconductor LDs have a typical emission pattern of 30°×30°; this provides for higher coupling efficiency than that of LEDs, particularly for single-mode fibers. LDs give peak power of -10 to +10 dBm, typically near 0 dBm, coupled into either multimode or single-mode fibers. The spectra of LDs are much narrower than for LEDs, ranging from 3 to 5 nm for typical multilongitudinal-mode lasers and 1 nm or less for single-longitudinal-mode lasers. The structure of a typical multimode laser is shown in Figure 2-28. The LD relies on mirrors formed by cleaved crystal facets at each end of the optical cavity. In this design, the 3- to 5-nm-wide emitted spectrum is made up of a group of lines with the lines separated by the mode spacing of the cavity. Efforts to narrow the spectrum to permit greater bandwidth and repeater spacing have resulted in other structural designs. As an example, the structure of an internal DFB laser to reduce the spectrum width to a few megahertz is also shown in Figure 2-28. Multimode laser operation produces mode-partition noise. This means that the relative intensity of the various modes fluctuates, while the total optical power remains constant. This fluctuation degrades system performance. Single-frequency lasers eliminate this noise. They reduce the dispersion bandwidth impairment to only that associated with the spectrum produced by the modulation. Further bandwidth impairments are caused by laser chirp, which is a frequency change in the laser output caused by on and completely off signaling. The change increases the undesired spectral lines. This effect can be suppressed by keeping the laser in the light-emitting state and changing the light intensity for signaling.

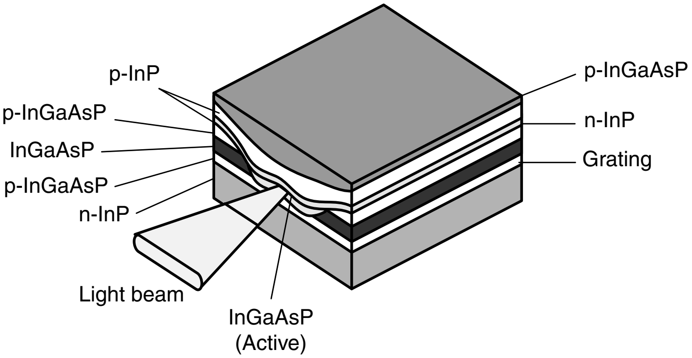
2.6.3 Line Signal

Optical line signals are usually in the form of return to zero (RZ) or nonreturn to zero (NRZ) pulses. The RZ pulses typically have an average 50 percent duty cycle. (Duty cycle is the fraction of time that the light source is on compared with the assigned time slot.) For off, the light from an LED is turned off; on the other hand, lasers are only dimmed (NRZ) to about 10 dB below the on light. Line signals are also used for clock recovery. The clock

Figure 2-28
LD structures.



(a) Cleaved crystal facets at each end of the optical cavity



(b) Grating for internal distributed feedback for narrowing of spectrum

recovery circuit requires a minimum number of transitions between 0s and 1s and only short runs of contiguously successive 1s or 0s. To meet these conditions, line signals are typically scrambled to provide a balance between 0s and 1s.

2.6.4 Tunable Lasers

With the advent of DWDM, the need arises to have tunable lasers that can be used in lieu of a large inventory of fixed-wavelength line cards. Today's DFB lasers are unique to a given wavelength; a DWDM system with, say, 100 channels presents a daunting challenge to the service provider in terms of sparing and maintenance. In case of an outage on a system, the carrier must be able to find the exact replacement laser. Tunable lasers also reduce system maintenance and troubleshooting activities, thereby reducing system operating and maintenance costs. A tunable laser is a single replacement for all static sources; sparing is done with a single, common, or configurable card. Tunable lasers are also at the basis of lambda bandwidth on demand and intelligent optical networking. Hence, transport nodes can match available capacity to real-time demand. Routes can be set up across the network for an end-to-end lambda circuit; this also enables the dynamic setup and teardown of lambdas based on downstream requirements. Proponents see these services entering the market on a wide scale in the near future. Also these devices support wavelength conversion, wavelength switching, and wavelength add/drop. The Intelligent Optical Network (ION), discussed in more detail in Chapter 7, “All-Optical Networks,” and Chapter 8, “GMPLS in Optical Networks,” is an architecture based on all-optical communication with a control mechanism for both the User-Network Interface (UNI) and the Network-Node Interface (NNI).

Tunable laser families include

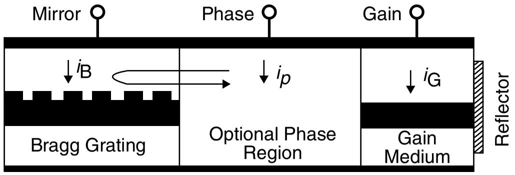
- Distributed Feedback (DFB)
- Distributed Bragg Reflector (DBR)
- Vertical Cavity Surface-Emitting Laser (VCSEL)
- External Cavity Diode Laser (ECDL)

Some aspects of this technology can be seen in Figure 2-29.[11]

LEDs, the least expensive fiber-optic transmitters, are used for multimode transmission at 850 or 1,310 nm, but they are limited in speed. FP LDs are relatively inexpensive and are typically used for single-mode transmission at 1,310 nm. DFB LDs are the most expensive and are used for single-mode transmission at 1,550 nm. The VCSEL is now popular for GbE operation at 850 nm. To give a rough view of the relative costs involved with these transmitters at the time of this writing, the Original Equipment Manufacturer (OEM) cost in volume of an LED transmitter/receiver (transceiver) component is about $20, an FP laser transceiver is about $100, a DFB laser is about $600, and a gigabit VCSEL is about $70.[6] The receivers in these types of relatively inexpensive transceivers are similar. They consist of a special diode that converts photons into electrons. Decent receivers are able to detect light levels down to hundreds of nanowatts. Note that the receiver does not care if the transmitter is multimode or single mode; it just knows that photons arrive, regardless of the path they take to get there.

Distributed Bragg Reflector Laser (DBR)

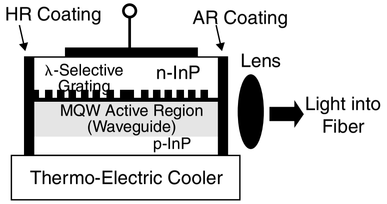
- Use periodic λ-selective Bragg gratings as reflector
- Tune by injecting current to vary the *n* of the Bragg grating and of the phase region → tuning range ~5 nm
- Current injection → tuning in ~nsec
- Variations include:
  - Sampled grating DBRs
  - Grating assisted codirectional coupler DBRs



Longer Phase Region → Wider Tunability
But More Losses → Lower Power

Distributed Feedback Laser (DFB)

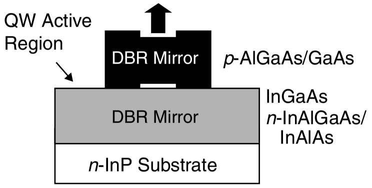
- Etch a "distributed" Bragg grating over active waveguide
- Stable, narrow single-mode operation for >2.5 Gb/s
- Power = 20 mW for signal; >200 mW for pumps
- λ-filter tuning achieved:
  - Thermally: ΔT of ~ 50°C → Δλ of ~ 5 nm
- Tuning range typically limited to 5~8 nm
- Multisection DFB have wider tuning but higher losses
- Fabrication of multisection DFBs can be complex



MQW: Multiple Quantum Wall

Vertical Cavity Surface-Emitting Laser (VCSEL)

- Mature technology for 850 nm (GaAs/AlGaAs)
- Tuning achieved via top mirror mounted on MEMS actuator
- Manufacturing advantages over edge-emitting lasers
- Easier packaging with vertical emitter and circular beam
- Lower power output than edge-emitting lasers [5-10 mW]



External Cavity Diode Laser (ECDL)

- Size of cavity set by external mechanical reflectors
- λ-filtering done external to laser diode. Tuning speed depends on type of actuation (Motor, MEMS, LC)
- High power and good spectral purity (such as Low-SMSR)
- Wide tuning range of > 100 nm limited by gain medium, not tunable filter

Example of a Littman-Metcalf ECDL

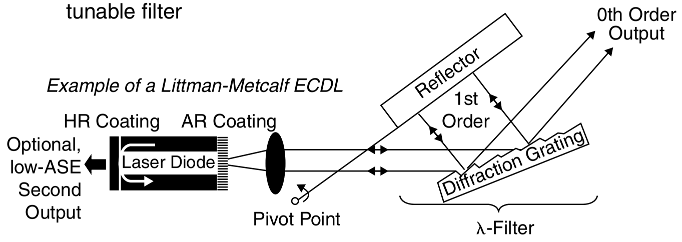


Figure 2-29
Tunable laser technologies.

2.7 Fiber-Optic Receivers

Optical receivers are coupled to the fiber strand as a unit. They convert on-off light pulses by photodetection into electrical signals. Those signals are processed by electronic components to provide the output data stream. There are two kinds of photodetector diodes: the Positive Intrinsic Negative (PIN)[12] and the avalanche photodetector diode (APD)[13] type. Typically, the APD type can detect lower optical signal power than the PIN type by about 10 dB (at 10^{-9} BER) (see Figure 2-30). An ideal receiver should offer good sensitivity (low received power for 10^{-9} BER), wide dynamic range, relatively low cost to manufacture and operate, reliability, low dependence on temperature, and so on. None of the diodes currently in production have all of the ideal characteristics. For example, the APD has the highest sensitivity and a good dynamic range, but it is difficult to produce. InGaAs photoconductors have poorer sensitivity, but are simpler and have the potential to be integrated with low-noise amplifiers. Over time, advances are expected to provide improved receivers with integrated optic and electronic circuits.

In summary, Table 2-5 recaps the key parameters associated with the optical transmission links we've just explored. Table 2-6 lays out some link engineering issues that affect systems at the current OC-48/STM-16 (approximately 2.5 Gbps) and OC-192/STM-64 (approximately 10 Gbps) rates.[14,15]

Figure 2-30
Minimum detectable power as a function of data rate for APD and PIN photodiodes.

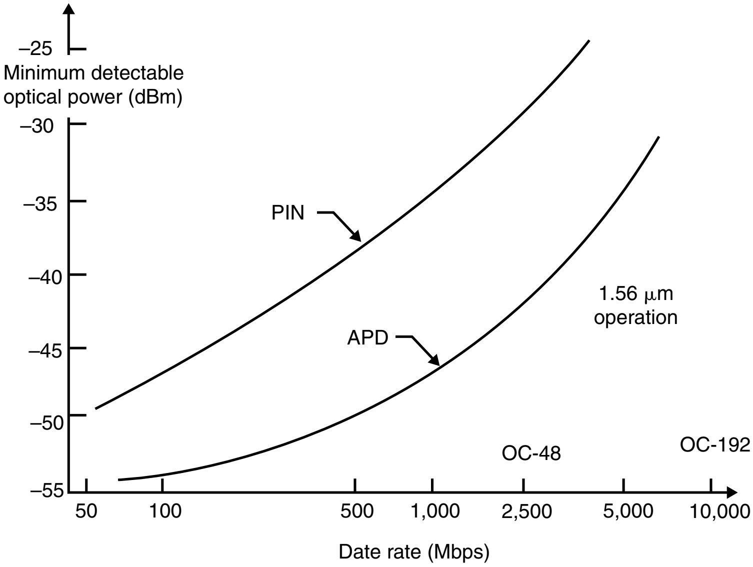


Table 2-5

Typical Optical Line Parameters

Wavelength	1.3 nm (nom.) or 1.5 nm (dispersion shifted)
Bit rate	2.5 Gbps (nom.); 10 Gbps being increasingly introduced
Code	Unipolar NRZ
Transmitter	Power (laser): -3.5 dBm Spectral width (half power): 0.5 nm
Receiver sensitivity	-32 dBm (for BER $= 10^{-11}$)
Power penalties	Dispersion: 0 dB (not limiting) Reflection: 0.5 dB
CO losses	2.5 dB
Loss margin	3.0 dB
Maximum outside plant losses	0.46 dB/km (0.74 dB/mile)
Maximum allowable dispersion	3.4 ps/nm/mile
Repeater spacing (long haul)	30 miles (more with lower losses)
WDM	16 (low-end), 32 (midrange), 80 (high-end) channels each at OC-192 (typical commercially available systems) with spacing in the 100-GHz, 50-GHz, and even the 25-GHz range

Table 2-6

Typical Link Engineering Considerations

OC-48/STM-16	OC-192/STM-64
■ Moderate OSNR requirement	■ Has 6 dB more OSNR than OC-48
■ Trend to dense channel spacing	■ Trend to dense channel spacing
■ Dispersion not a major consideration ▪ SMF-28 requires narrow line-width transmitters	■ Dispersion a major consideration ▪ Requires narrow line-width transmitters ▪ Prechirping transmitters an option ▪ SMF-28 and some NZDSF links require dispersion compensation
■ PMD not an issue	■ PMD an issue over older fiber
■ SMF-28 ▪ Limited by OSNR	■ SMF-28 ▪ Limited by OSNR, dispersion, and PMD
■ NZDSF ▪ Balance between OSNR and FWM	■ NZDSF ▪ Balance between OSNR and nonlinearities
■ Typical specs ▪ 6×25 dB ▪ 8×20 dB	■ Typical specs ▪ 5×20 dB NZDSF ▪ 5×25 dB SMF-28 with dispersion compensation ▪ Next-gen Raman systems extend distance

2.8 Other Technologies

This section discusses other optical technologies that play a role in metro access and metro core spaces, as well as in the long haul.

SONET/SDH is an interworking standard developed in the late 1980s and early 1990s to define data rates, optical characteristics, and operations protocols over a data communications channel for support of multivendor optical transmission. This technology is explored in Chapter 3, "Traditional SONET Architecture," but common data rates are shown in Table 2-7.

WDM is a multiplexing technology. Optical multiplexers aggregate multiple lambdas onto a single waveguide (in this case, the fiber). Demultiplexers separate the combined signal into individual lambdas and output ports/fibers. Three basic categories of multiplexers are star couplers, thin-film filters, and phased array (see Figure 2-31). A star coupler is an all-fiber technology that is inexpensive and robust, but suffers from high signal loss. Typically, it is used for low-count systems. Thin-film filters are manufactured according to well-known semiconductor processes. They are a robust and proven technology that enables the network designer to pay as one scales. This technology can be used up to about 40 lambdas. A planar array waveguide is a silicon semiconductor technology that is robust and proven. This technology can be used for midrange systems (16 to 40 lambdas); it is cheaper on a per-channel basis than thin-film filters, but users must pay for the entire capability (the wafer) up front. Other technologies used in WDM include bulk gratings, fiber Bragg gratings (FBGs), Mach-Zehnder structures, and a combination of Mach-Zehnder structures and FBGs. (Thin-film filters, FBGs, and array waveguides are discussed in greater detail in the following section.)

Table 2-7

Common Data Rates

Siglum	**Data Rate (Mbps)**
OC-1/STS-1	51.84
OC-3/STS-3/STM-1	155.52
OC-12/STS-12/STM-4	622.08
OC-48/STS-48/STM-16	2,488.32
OC-192/STS-192/STM-64	9,953.28
OC-768/STS-768/STM-256	39,813.12

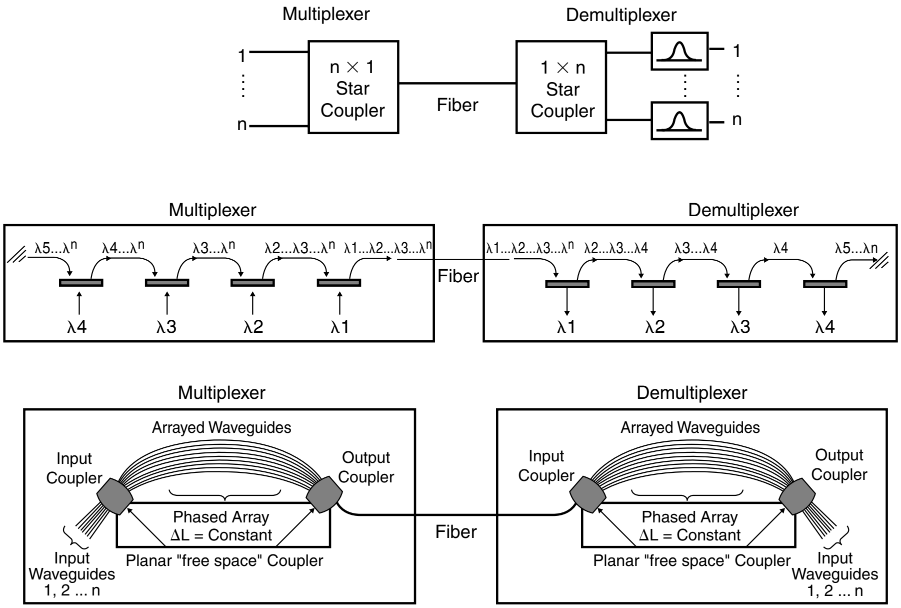


Figure 2-31 *Multiplexing technologies. Top: Star coupler. Center: Thin film. Bottom: Arrayed waveguide.*

WDM can be viewed as a sublayer to SONET. WDM can replace a number of distinct SONET transmission systems with a single WDM transmission system and a single-fiber pair that support the (numerous) constituent SONET end systems, as shown in Figure 2-32. DWDM systems saw major deployment in the United States in the late 1990s.[16] Figure 2-33 shows the optical window of operation for WDM. Table 2-8 provides a view of the system-level evolutions over time. Table 2-9 reviews the evolution of all-optical networks, whereas Figure 2-34 shows technology generations directly applicable to WDM. Finally, Table 2-10 summarizes the entire space of optical technologies on a time line. Using FDM techniques on a fiber transmission system, pair gain was achieved where multiple signals operating at different frequencies could each carry OC-48 or OC-192 channels between endpoints. Signal spacing improved to the 100-, 50-, and even 25-GHz range.

In long-haul applications, optical amplification conjoined with WDM enables the network planner either to space out hub, regeneration, or amplification sites (as shown in Figure 2-35) or to skip amplification at some sites using techniques such as Raman amplification (as shown in Figure 2-36). Figure 2-37 shows another view in the trend towards all-optical networks.[17]

Figure 2-32
DWDM designs.

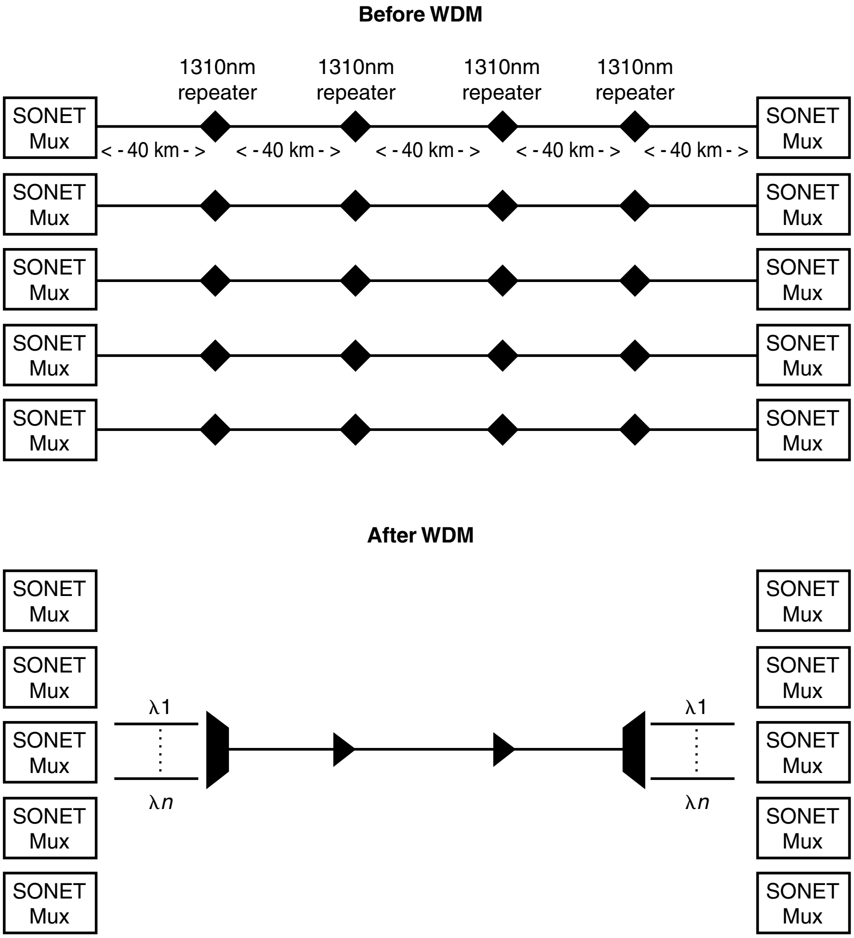


Table 2-8

System-Level Evolutions, 1977—2007

Phase 1 (1977–present)	Asynchronous TDM (optical DS3s and so on)
Phase 2 (1989–present)	Synchronous TDM (SONET/SDH)
Phase 3 (1990–present)	WDM
Phase 4 (1994–present)	DWDM
Phase 5 (2001–present)	ION/optical time-division multiplexing (OTDM)

Figure 2-33
Window of operation for DWDM.

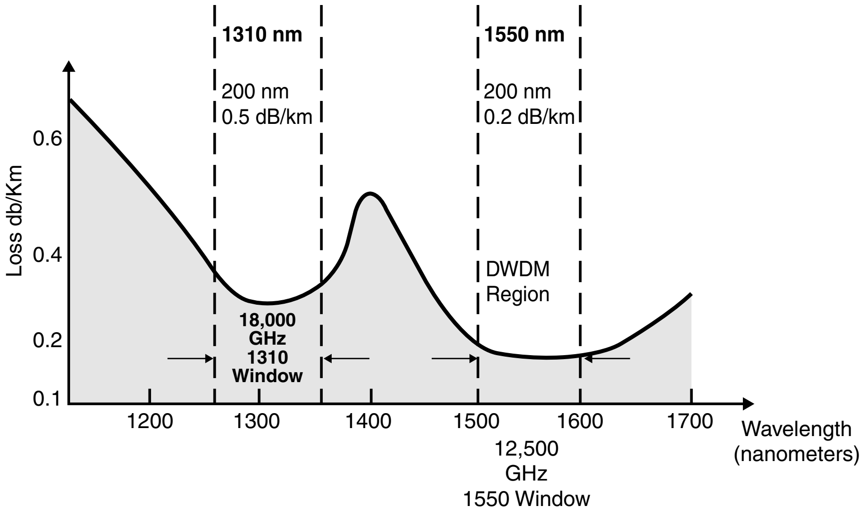
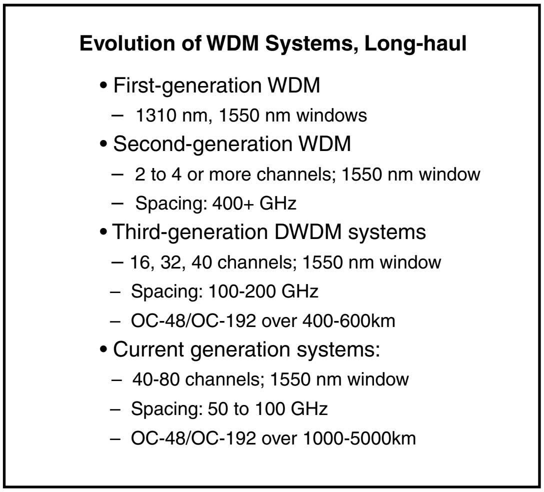


Table 2-9

Evolution of All-Optical Systems, 1990–2000

	Early 1990s	Mid-1990s	Late 1990s	Early 2000	Mid-2000
Mesh networks			100-GHz WDM, 50-GHz WDM, Static OADMs	Dynamic OADM	Dynamic optical switches, wavelength changers
Ring networks		200-GHz WDM	100-GHz WDM, 50-GHz WDM, Static OADMs	Dynamic OADM	Dynamic optical switches, wavelength changers
Linear networks	400-GHz WDM	200-GHz WDM	100-GHz WDM, 50-GHz WDM, Static OADMs	Dynamic OADM	Dynamic optical switches, wavelength changers



Evolution of WDM Systems, Long-haul

- First-generation WDM
  - 1310 nm, 1550 nm windows
- Second-generation WDM
  - 2 to 4 or more channels; 1550 nm window
  - Spacing: 400+ GHz
- Third-generation DWDM systems
  - 16, 32, 40 channels; 1550 nm window
  - Spacing: 100-200 GHz
  - OC-48/OC-192 over 400-600km
- Current generation systems:
  - 40-80 channels; 1550 nm window
  - Spacing: 50 to 100 GHz
  - OC-48/OC-192 over 1000-5000km

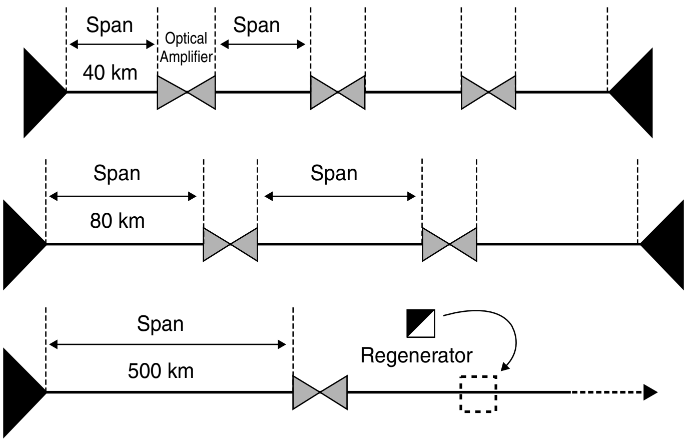
Figure 2-34 Evolution of WDM systems (long haul).

Table 2-10 Evolution of Optical Technologies, 1996–2010

Phase	Phase 1	Phase 2	Phase 3
Time frame	1996–2000	2001–2005	2006–2010
Major application	Voice	Voice + data	Voice + data + optical apps
Major architecture	SONET	SONET + ION	SONET + ION + packet optical
Characteristic	Static	Wavelength routing	Optical processing
Technologies	Thin-film filters, FBGs, arrayed waveguides, optical amplifiers/ EDFA, lasers, receivers, FEC	Optical switching, MEMS (2D and 3D), bubble jet, tunable lasers, dynamic gain flattening, dynamic OADM, Raman amplifiers, tunable dispersion compensation, some General MPLS (GMPLS) deployment, OC-768	OTDM, optical 3R, lambda conversion, optical FEC, optical switching, 40- or 100-Gbps Ethernet, some FTTH

Figure 2-35 Spacing out of hub, regeneration, or amplification sites (long haul).

Hut Spacing

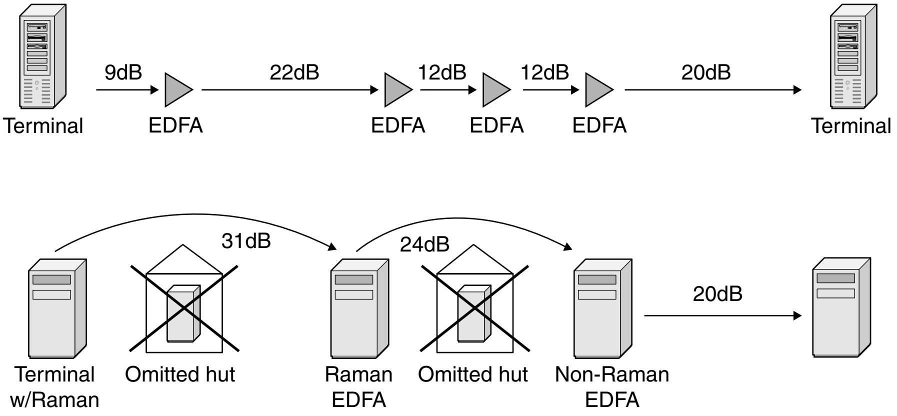


- Short hut spacing
  - – 40 - 80 km, 10 - 20 dB loss
  - – Advantageous to high-speed signals, supports many waves
  - – Less regenerators, but requires more huts and more amplifiers
- Long hut spacing
  - – 80 - 120 km, 20 - 30 dB loss
  - – Reduces first costs, but is less advantageous to transmission

Figure 2-36 Elimination of hub, regeneration, or amplification sites (long haul).

Advanced intelligent optical transport systems with optional Raman amplification result in equipment cost savings by eliminating short-span repeater sites

- Lower capital investment
- Reduced space, power
- Lower operational costs – fewer sites to maintain



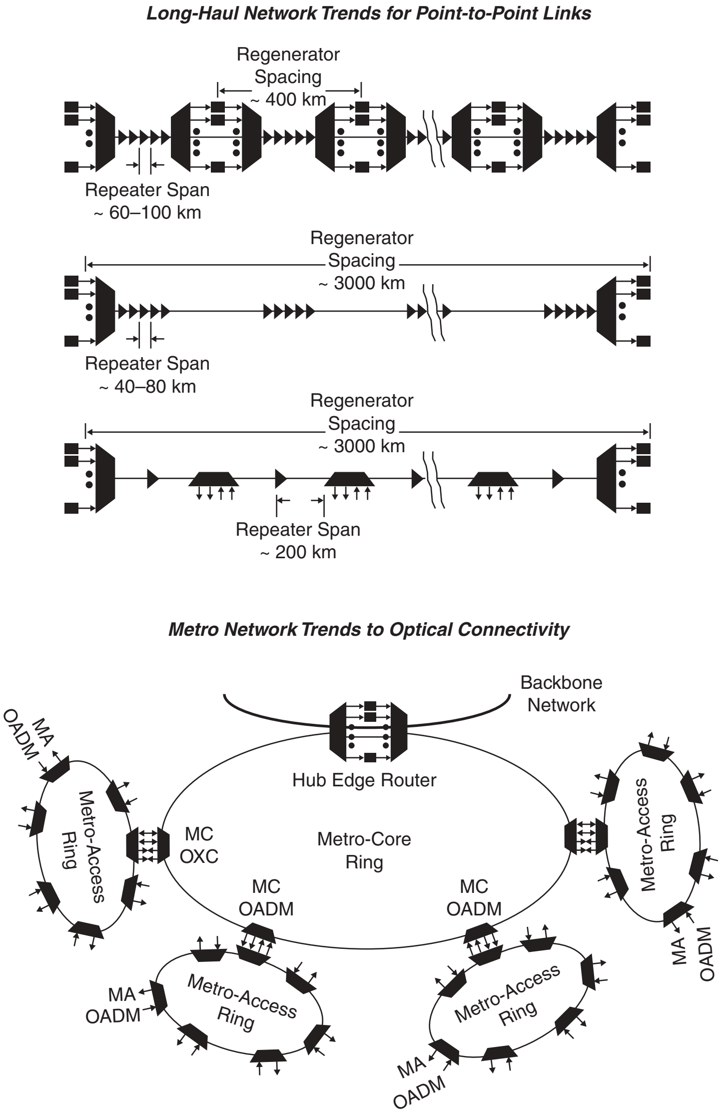


Figure 2-37
Trends in regeneration and the transition to optical networking.

2.9 DWDM Enablers

2.9.1 Thin-Film Filters

Thin-film filters can be used to demultiplex wavelengths in a WDM system (see Figure 2-38). Thin-film filters rely on an interference effect created by a highly reflective mirror and a linear variable filter that are coated onto opposite sides of a glass substrate. Typically, several wavelengths are passed through the filter on a single beam of light; as the beam of light encounters the variable filter, the demultiplexing function occurs. The process can also be designed to perform the multiplexing functions. Thin-film filters offer a cost-effective, high-performance solution for low to medium channel counts, and unsurprisingly thin-film technology is the primary method used for networks of 16 to 40 channels with 400-GHz and/or 200-GHz spacing.

The perennial manufacturing goal is to continuously improve product yields; add to that the goal of finding economically feasible methods for producing 100- and 50-GHz channel spacing or channel counts. A solution in support of the goal of narrower channel spacing and higher channel counts is interleaving technology. Figure 2-39 shows how interleaving technology can be used to achieve 50-GHz channel spacing. On the left-hand side of the figure, interleaving combines two separate multiplexing and demultiplexing devices. One mux interleaves the odd channels, whereas the second interleaves the even channels. On the right-hand side of the figure, 200-

Figure 2-38
Thin-film filters.

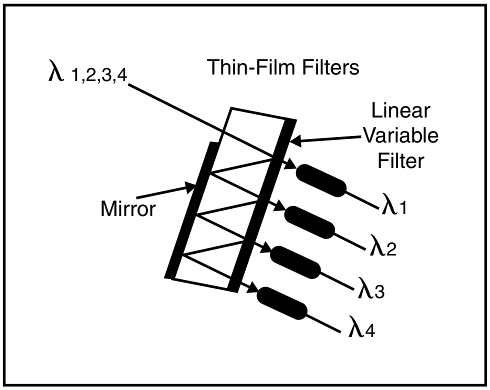
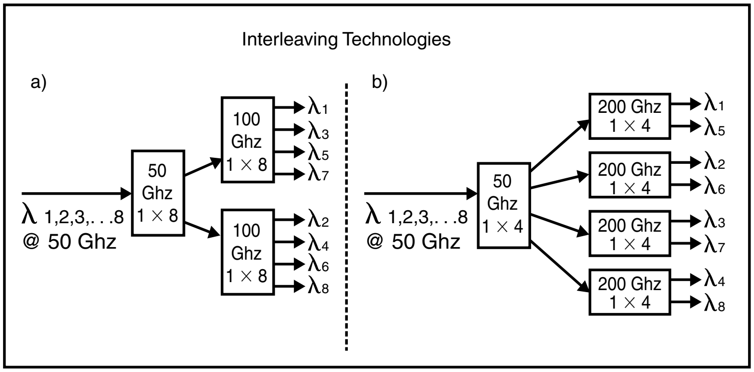


Figure 2-39
Interleaving technologies.



GHz spacing is interleaved to support 50-GHz applications. This approach adds channels in groups of four instead of two. One benefit of interleaving is that it enables a carrier to add capacity on an as-needed basis. Remember that this approach tends to become costly in the higher channel count systems. In addition, as the channel count increases beyond 16, insertion loss increases in the overall system.

2.9.2 Fiber Bragg Grating (FBG)

FBG is a technique used to separate and combine wavelengths. It does so by changing the structural core of standard single-mode fiber. The fiber strand is exposed to a high-power ultraviolet light, which creates a periodic change in the refractive index along the core by creating ridges on a short section of fiber. The periodic change in the fiber core enables narrowband wavelength selection. Each fiber grating isolates a single wavelength (see Figure 2-40). Advantages of FBG technology include wavelength uniformity and low insertion loss. The primary limitation is that it performs better at moderate channel counts (under 16 channels), and it is difficult to scale for the higher channel count systems.

2.9.3 Arrayed Waveguide Grating (AWG)

AWGs are painstakingly built by depositing carefully engineered and composed glass layers on a silica or silicon substrate. To guide light pulses through the gratings, glass layers are patterned and etched with standard semiconductor processes. Figure 2-41 illustrates how the input signal,

Figure 2-40
FBG.

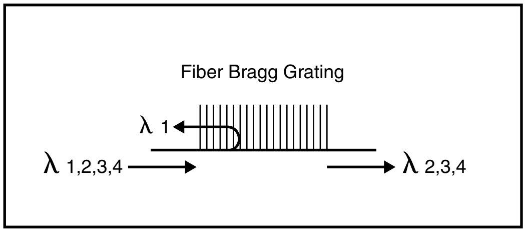
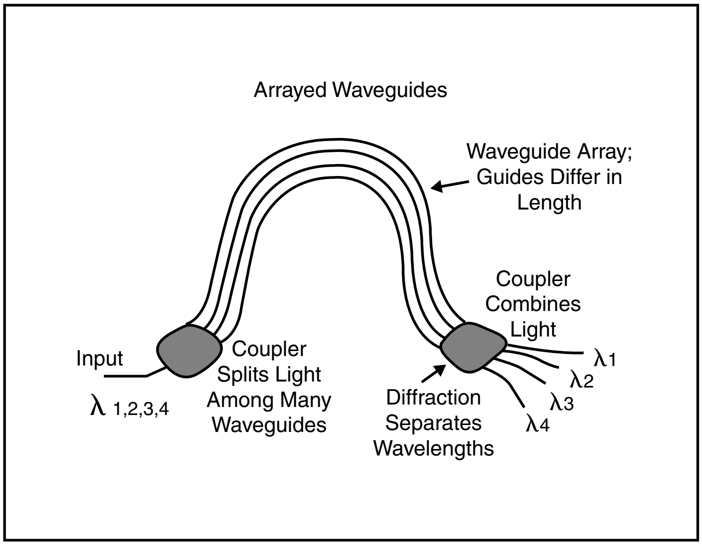


Figure 2-41
Arrayed waveguides.



arriving from a distant destination over a single fiber, is fractioned by the origination coupler and passed on to multiple curved waveguides. Each waveguide is a different length; this enables the light signals to travel at different speeds across the AWG. This difference in speed gives rise to interference and diffraction when the signals enter the receive coupler. Diffraction separates the wavelengths and enables the individual light signals to exit the AWG at the same time, but on different fibers. These fibers are then connected to receive-SONET terminals. As you might expect, the opposite process maps the signal from multiple transmit-SONET terminals onto a single fiber.

Using AWGs as an alternative to thin-film filters and FBGs reduces the overall cost and insertion loss of the network. Therefore, AWGs have become the technology of choice for high-channel-count DWDM systems. Only one grating is needed to meet the channel count and spacing requirement of the network, whereas thin-film filters require interleaving technology and additional filters to meet higher-channel-count needs.

2.9.4 Regenerators and Optical Amplifiers

As we've mentioned repeatedly, the signal attenuates as it travels through the fiber waveguide. Mechanisms to restore the signal to its pristine condition must be put in place, particularly for long-haul applications. The following list defines key elements in this process. Figure 2-42 displays the key functions of a regenerator.

SONET regenerator

- a) Full duplex (dual simplex).
- b) Operates on one wavelength in each direction.
- c) Input wavelength equals the output wavelength in each direction.
- d) Forward wavelength may or may not equal the reverse wavelength.
- e) Couples SONET Sections within a Line.
- f) Bufferless. Pipeline fixed delay. Transmit clock is the recovered receive clock.
- g) Examines and/or writes SONET section overhead.
- h) Payload, Line, and Path information are passed through unmodified.
- i) Unscrambles and rescrambles to support item g.
- j) Keeps output Section active even if input Section fails (fault isolation).
- k) Reamplify, reshape, and retime (3R function).
- l) Comprises section-terminating equipment (STE).

Wavelength Regenerators

- Performs "3R" signal processing (retime, reshape, regenerate) through optical-to-electrical-to-optical (O-E-O) conversions
- Removes accumulated signal noise

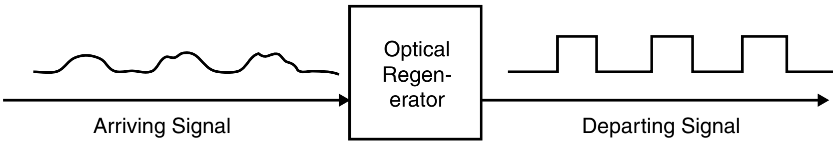


Figure 2-42 Wavelength regenerators.

SONET Line-terminating equipment (LTE)

- a) Includes the functions of a SONET regenerator.
- b) Terminates a Line and the associated Section.
- c) Operates on Stratum clock.
- d) Multiplexes Paths onto a Line and demultiplexes a Line into Paths.
- e) It is a dual-simplex Path mux.
- f) Supports maintenance/protection switching for muxed Paths between LTEs.

SONET Path-terminating equipment (PTE)

- a) Terminates a Path and the associated Line and Section.
- b) Operates on local clock.
- c) Equipment is where the SONET path overhead is processed.

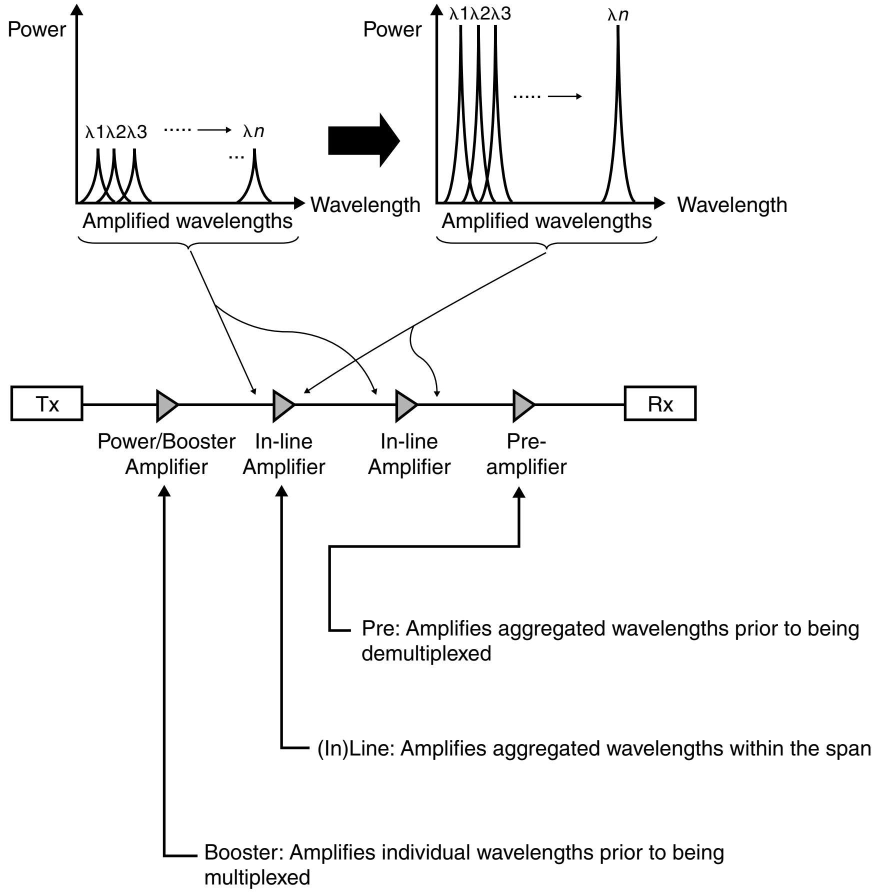
SONET transponder

- a) Couples differing optical physical-medium-dependent functions back to back (for example, wavelength 1 to wavelength 2, multimode to single-mode, 850 nm to 1,300 nm).
- b) Dual simplex.
  - Passive transponder.
- c) Transparently passes all bits (items a and b).
  - Active transponder.
- d) A special case of SONET regenerator supporting item a.
- e) Terminates Section (see regenerator) and Line (see LTE).

Preferably, the amplification function in an optical domain can be handled without resorting to the optical-electrical-optical (O-E-O) conversions used by devices in the previous list. The functionality of a general optical amplifier is shown in Figure 2-43.

An optical amplifier, such as simple EDFAs in Figures 2-44 and 2-45, is used to boost an attenuated input signal. The amplifier is the repeater technology that is used in long-haul applications, oceanic cables, and some metro core applications. A basic EDFA consists of isolators, a pump laser, an erbium-doped fiber, and a coupler. The isolators reside at the front and back ends of the amplifier to prevent pump-laser emission or suppress light reflection coming from the erbium-doped fiber. The pump laser supplies power to the fiber carrying the light signal, and a coupler connects the pump laser directly to the amplifier. In simple EDFAs, amplification takes place in 50 m or so of erbium fiber coiled inside. Erbium in the doped fiber

Figure 2-43 Functions of a general optical amplifier.



Advantages over Electronic Systems
- No OE, EO conversion
- Low noise, high gain
- Greater bandwidth than electronic repeaters
- Transparent to bit rates
- Transparent to modulation formats
- Simultaneous generation of multiple WDM wavelengths

provides maximum amplification at 1,550 nm. Figure 2-46 is a typical amplification profile.

A gain-flattening filter can extend an EDFA amplifier's range of operation, as shown in Figure 2-47.

Introducing optical amplification into a long-haul link (see Figure 2-48) obviates the need for retiming, reshaping, and regenerating (otherwise known as 3R). However, EDFAs have a tendency to create noise in the

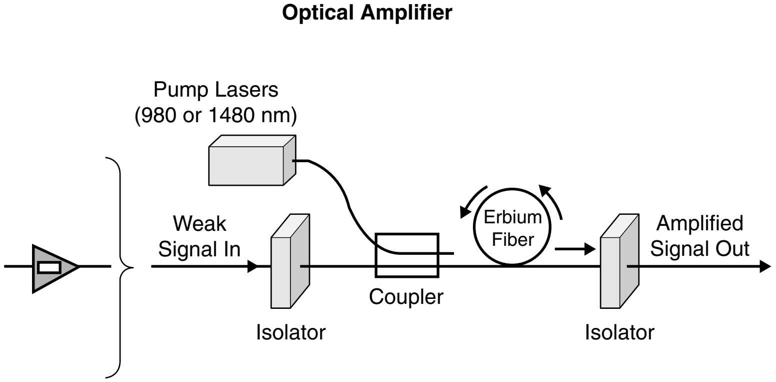
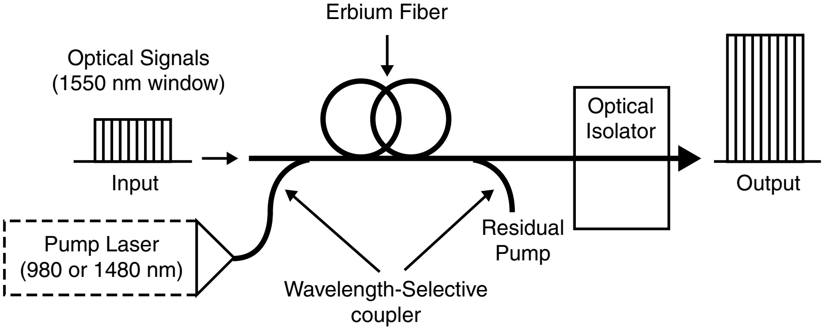


Figure 2-44
Optical amplifier.

Figure 2-45
EDFA structure.

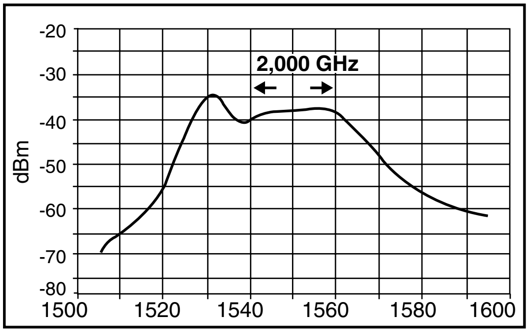
Erbium-Doped Amplifier (EDFA) Structure

- Simultaneously amplifies wavelengths in the 1530-1565 nm region
- Inherent spontaneous emissions are also amplified and become noise



fiber-optic network, and electrical regeneration is still needed at various points in the network to suppress that noise. Regeneration, we know, is expensive and limits the bandwidth that can be supported. To regenerate an optical signal, a wavelength must be converted from an optical signal to an electronic signal and back to an optical signal (O-E-O conversion). Hence, in a typical DWDM network, each optical wavelength traveling through a fiber has to be routed through an expensive SONET regenerator that performs 3R functions. Although performance varies by manufacturer, wavelengths traveling at higher signal rates (for example, OC-192) could need up to eight full bays of SONET equipment to regenerate just eight wavelength signals.

Figure 2-46
Typical EDFA amplifier profile.



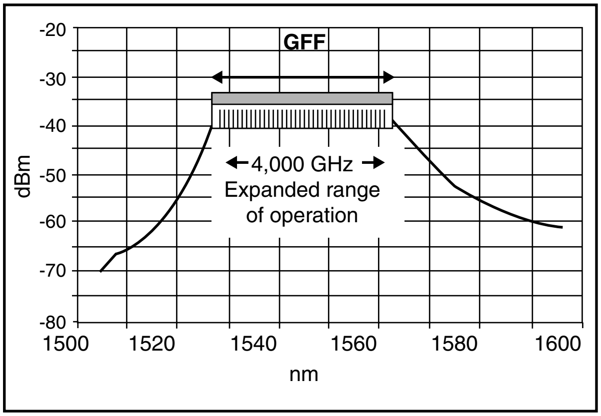
On 2,000 GHz window:
5λs when spaced 400 GHz
10λs when spaced 200 GHz
20λs when spaced 100 GHz
40λs when spaced 50 GHz
80λs when spaced 25 GHz

EDFA

f @ 1530 nm = 1.960×10^{14}
f @ 1540 nm = 1.948×10^{14}
f @ 1550 nm = 1.935×10^{14}
f @ 1560 nm = 1.923×10^{14}

Figure 2-47
Gain-flattening filter advantages.

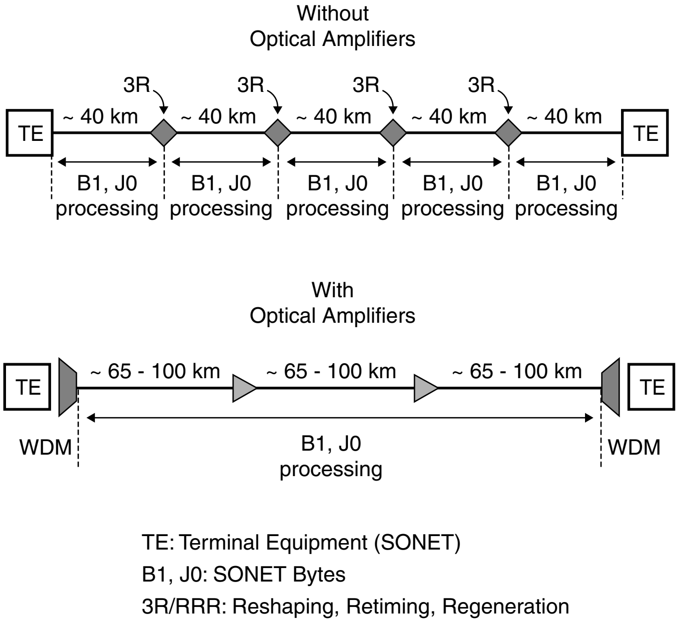
Gain-Flattening Filter (GFF) Advantages



On 4,000 GHz window:
10λs when spaced 400 GHz
20λs when spaced 200 GHz
40λs when spaced 100 GHz
80λs when spaced 50 GHz
160λs when spaced 25 GHz

Raman amplification, a new technology for next-generation DWDM systems, mitigates the costly regeneration process. It operates by enabling wavelengths to travel further along the fiber between amplification points. Whereas EDFA amplification takes place in the 50 m of fiber coiled inside the box, Raman amplification takes place outside the box, along the regular fiber in the trench, for about 10 km or so. This improves the performance of amplification and reduces the number of regenerators the network needs.

Figure 2-48 Reduction of regeneration when using optical amplifiers.



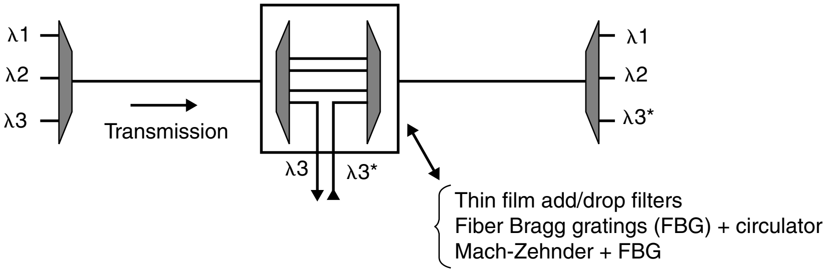
2.9.5 Optical Add/Drop Modules (OADMs)

Multiplexers called optical add/drop modules (OADMs) are used to deliver signals to their proper destinations. OADMs can add or drop an individual wavelength or a set of wavelengths without electronic conversion. They can consist of a preset fixed wavelength or be dynamically configurable for one or more wavelengths (see Figure 2-49).

The filtering effect is achieved via a number of technologies, notably thin-film and FBG technology, as shown in Figure 2-50. One factor most often associated with OADMs, however, is loss. Circulators access or transport (add or drop) signals of different wavelengths traveling on the fiber; these are passive devices that guide wavelengths from port to port in one direction only. The purpose of circulators is to guide the wavelengths that are filtered in and out of the OADM without interrupting any of the light signals. For example, in Figure 2-51, the input light signal is carrying multiple ($\lambda_1, \lambda_2, \lambda_3, \ldots \lambda_x$) channels. As it moves through the OADM, FBG technology and the circulators enable channel λ_x to be dropped and channel λ_y to be added.

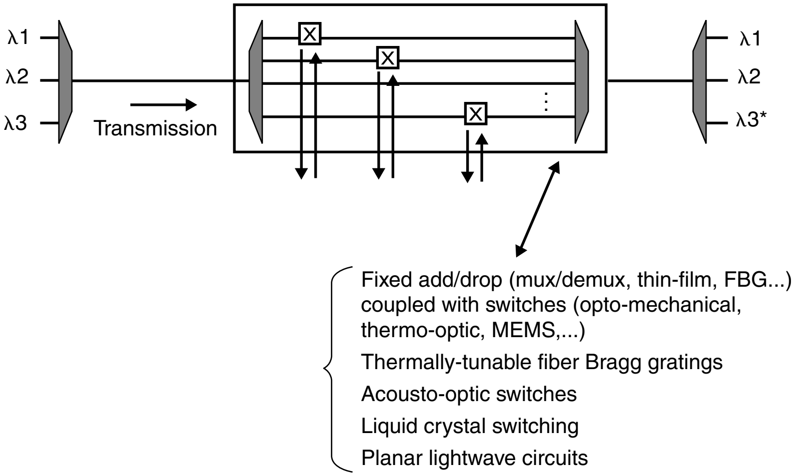
OADMs bring planners a step closer to the all-optical network. They're useful network elements in the context of metro networks. OADMs placed

Figure 2-49
Fixed/configurable OADMs.



Thin film add/drop filters
Fiber Bragg gratings (FBG) + circulator
Mach-Zehnder + FBG

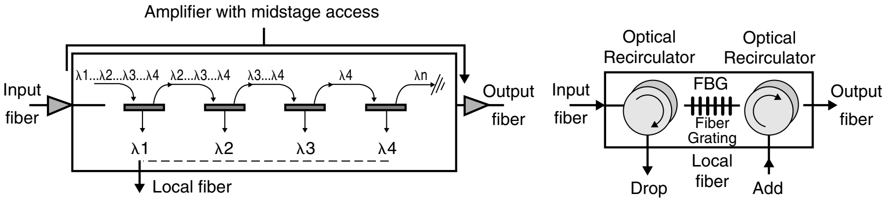
(a) Fixed-**λ** OADM



Fixed add/drop (mux/demux, thin-film, FBG...) coupled with switches (opto-mechanical, thermo-optic, MEMS,...)
Thermally-tunable fiber Bragg gratings
Acousto-optic switches
Liquid crystal switching
Planar lightwave circuits

(b) Dynamic (Configurable) OADM

Figure 2-50
OADM technologies (example).



- Low loss
- Service disruptive when adding/changing λ

- Higher loss for low channel counts
- Service disruptive when adding/changing λ

(a) Fixed thin-film OADM

(b) Fiber gratings OADM

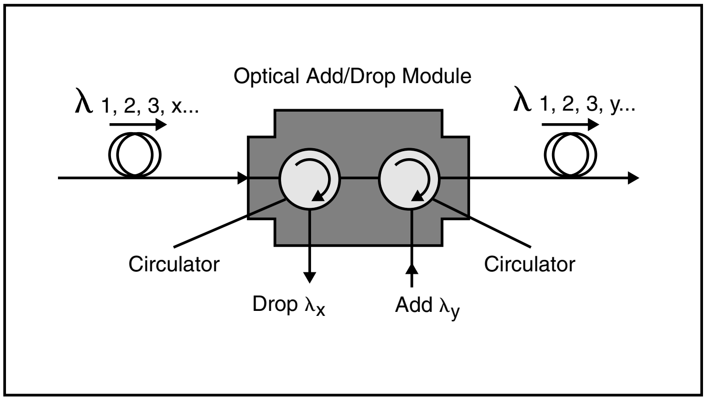


Figure 2-51
OADM.

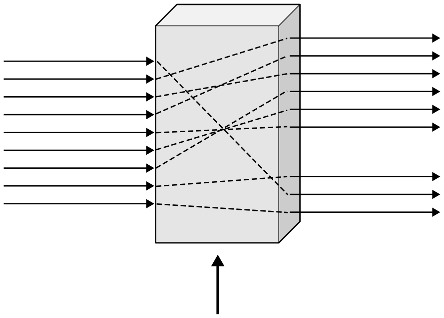
between two end terminals along any route can replace an optical amplifier. Commercially available modules enable carriers to drop or add up to four STM-16/OC-48 channels between DWDM terminals. OADMs enable certain wavelengths to pass through the node uninterrupted and also have broadcast capabilities for dropping information on several channels (typically four) to become express channels. Optical add/drop is well suited for meshed or branched network configurations as well as for ring architectures used to enhance survivability.

OADMs are reasonably inexpensive and, in particular, are cheaper than electronic ADMs. Fixed-lambda systems are mature and widely deployed; configurable-lambda systems are maturing rapidly. Configurable-lambda systems are tapped for use in the IONs discussed in Chapter 7.

2.9.6 Optical Crossconnects (OXCs)

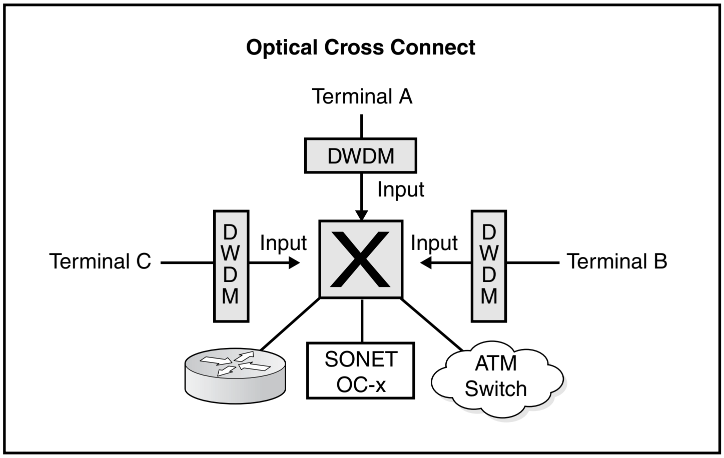
Many current-generation optical networks use OXCs. Figure 2-52 depicts OXC system functionality. OXCs are capable of interconnecting multiple DWDM systems (for example, Terminals A, B, and C in Figure 2-53) and switching optical channels in and out of DWDMs without having to convert the signal into a lower-speed electrical signal. OXCs also facilitate interconnections between the DWDM systems and Internet Protocol (IP) routers, SONET/SDH multiplexers, and Asynchronous Transfer Mode (ATM) switches. Their deployment is critical to all-optical networks, not just because they support higher bandwidth, but because they also support more sophisticated network designs. For example, OXCs can perform traditional SONET/SDH functions (such as restoration, signaling, perfor-

Figure 2-52
OXC function.



- Micro-electrical systems (2D and 3D MEMs)
- Thermo-optic
- Opto-mechanical
- Bubble jet

Figure 2-53
OXC example.



mance monitoring, fault management, and multiplexing). Putting those functions in the optical layer gives carriers a flexible way to reconfigure their networks for the same or even lower deployment and operation costs. The optical core uses OXCs to manage high-speed traffic, while traditional SONET equipment moves to the edge to manage the lower-speed interfaces and protocol mixes. Also, non-SONET equipment such as GbE switches, 10 GbE switches, and Resilient Packet Ring (RPR) devices can be supported by the optical core.

First-generation (1G) crossconnects had optical inputs and outputs, but used an electronic switching matrix in between. Here, in order to switch the wavelengths, the crossconnect converts optical signals to electrical signals and then converts them back to optical signals. This O-E-O conversion makes the switching process inefficient and expensive. Second-generation (2G) crossconnects have the capability of switching wavelengths without an electrical conversion (O-O).

There are a number of advantages to using OXCs:

- All-optical crossconnects have a simpler design than O-E-O devices. The number of devices and interfaces used for O-E-O conversion is eliminated from the all-optical crossconnect. This speeds up network upgrades because various elements and interfaces do not have to be adjusted to accommodate new speeds and feeds of the system.
- It is easier to operate the network with an O-O device because of the automation and remote control functions available.
- Testing and provisioning can be performed on the entire network from a single location.

Three primary technologies have been developed to build all-optical crossconnects:

1. Micro-Electro Mechanical Systems (MEMS)
2. Bubble technology
3. Liquid crystal technology

Micro-Electronic Mechanical Systems (MEMS) MEMS technology is used in wavelength switching and OXCs. MEMS devices are based on the mechanical positioning of very small mirrors that direct wavelength signals between input and output ports. These compact devices are silicon wafers with a matrix of moveable mirrors. When switching output ports, the mirror at the intersection of output 1 and output 2 rises and redirects the light signal coming from input 1, while the other mirrors on the matrix remain flat. MEMS devices are attractive because of their miniaturization and potential mass production. The main drawback of an MEMS device is the issue of movable parts and reliability.

Bubble Technology With bubble technology, an input signal enters the device with the goal of exiting at output 1. When an input signal and output signal intersect in the device, a vapor bubble is formed upon command. This point of intersection houses an index-matching liquid that enables the formation of the bubble. In the presence of a bubble, the input signal is reflected toward output 2. (Without the formation of a bubble, the light signal passes straight through the device toward output 1.) The bubble is of interest because of its connection with a proven technology and lack of moving parts. However, scaling the system can become a problem. Each time the service provider adds a block, more insertion loss is generated and accumulated across the system.

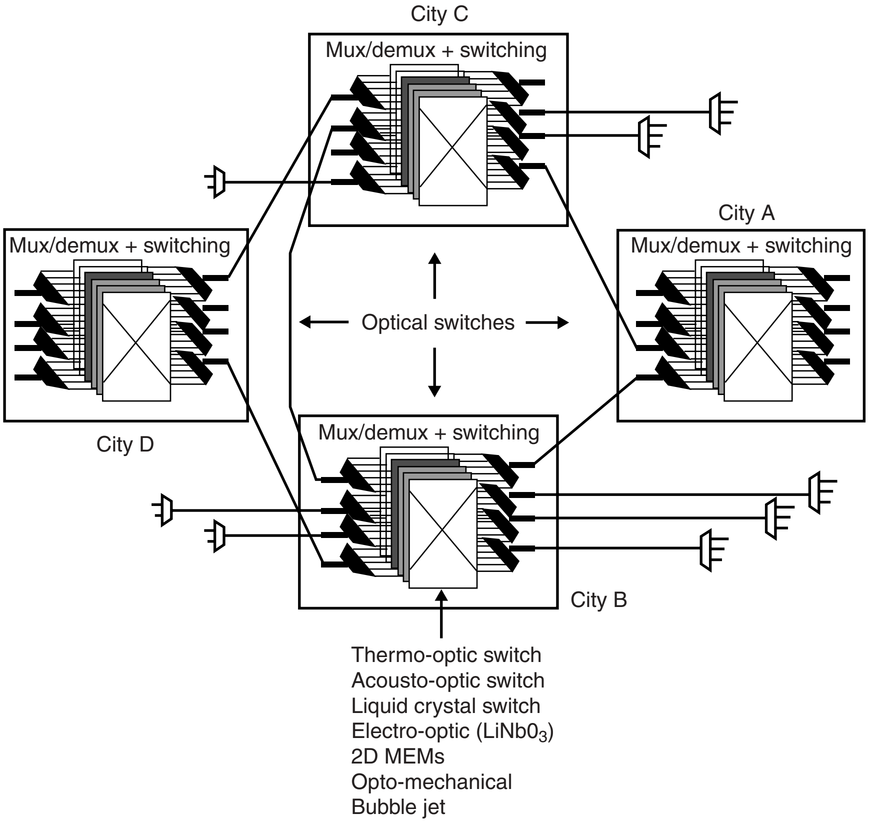
Liquid Crystal Technology Liquid crystal displays (LCDs) consist of active and passive elements. The active element is the liquid crystal cell and the passive element is the polarization beam splitter. The switching action is dependent on whether or not an electrical field is present. When there is no electrical field in the device, the input signal reflects off the liquid crystal pixel and rotates 90 degrees. When an electrical field is present, the light signal travels straight through the active cell without changing direction. Two advantages of LCD technology are the lack of movable parts, making it more reliable than MEMS, and the manufacturing process (interestingly, the same used for flat panel displays), which uses a technology that is mature so that materials, fabrication, infrastructure, and processes are already in place and scaled for large quantities. The drawbacks of liquid crystal technology are thermal sensitivity, slower switching speeds, and limited scalability. Thermal sensitivity can be overcome with temperature controllers, better design, and improved liquid crystal fluid, but scalability is always an issue because of cost.

2.9.7 Optical Switching

Optical switching is the enabling technology that makes ION possible by enabling the direct switching of constituent wavelengths, whether intrinsic to the network or customer originated. Optical switching eliminates O-E-O conversions. Figure 2-54 depicts a network application example. Like any other switching system, optical switches comprise a switching matrix and a switch control function. (We'll look at the control function again in Chapters 7 and 8.)

Optical-switching matrices in the context of metro networking need to fulfill certain requirements. They must have proper optical characteristics such as low insertion loss, high isolation, and low back reflection. In addition, they need to be highly reliable, fast, compact, low power consuming, and scalable. It is also important that the resulting switch is integrated into an end-to-end management system. Techniques that are suitable for optical-switching matrices include but are not restricted to MEMS, some

Figure 2-54 Network of optical switches.



MEMS-like technologies (thermo-capillary and piezo-electric), liquid crystals (LXT), bubbles, thermo- and acousto-optics, and holograms. These technologies offer different characteristics, which are summarized in Table 2-11.[18]

MEMS usually refers to arrays of tiny tilting mirrors sculpted from semiconductor materials such as silicon. In 3D arrays, the mirrors can be tilted in any direction. 3D MEMS arrays can be used for large-scale OXCs, particularly if groups of wavelengths can be switched from one fiber to another. In 2D arrays, the mirrors can only flap up and down incorporating a single axis. 2D MEMS arrays will be used in smaller, 1G, all-optical switches because they are further along in the development process. As MEMS is based on mechanical movement, it offers the best isolation performance available. For the same reason, MEMS cannot be used for drop-and-continue applications.

Table 2-11

Characteristics of Switching Technologies

	MEMS	LXT	Bubbles	Thermo-optics	Acousto-optics	Holo-grams
Scalability	Very high	Medium	Medium	Medium	Medium	High
Switching speed	Medium	Medium	Medium	Medium	Fast	Very fast
Reliability	Moving parts	Good	Good	Good	Good	Good
Loss	4×4: 3 dB 16×16: 7 dB	2.5–6 dB, up to 32×32	32×32: 5 dB	8×8: 8 dB	1×2: 2.5 dB	Low
Power consumption	Medium	Very low	Unknown	Very high when based on silica	Unknown	Medium, but requires high voltages

In LXT, light is passed through liquid crystals that alter their transmission characteristics when an electric current is applied. Most LXT devices depend on the polarization properties of the light for switching. LXT offers particularly low power consumption. LXT devices can throttle the amount of light that passes through them. They can therefore be used as variable attenuators or power splitters with a configurable splitting ratio. The latter enables various drop-and-continue applications. LXT switches have good scalability; however, the technology will first be used for modestly sized, wavelength-selective devices. LXT can be combined with electroholographic technology for increasing switching speeds.

As discussed previously, bubbles formed by printer ink pens have surfaces that act like mirrors. Bubbles are currently used for optical switches by reusing parts of the Desk Jet printer technology. The technology is intended for use in add/drop multiplexers (ADMs) and scalable switches.

Thermo-optics consists of passive splitters. When these are heated, their refractive index is changed to alter the way in which they divide wavelengths between either of the two output ports. They are wavelength sensitive and need temperature control. Thermo-optics are ideally suited to integration. AWG can be integrated with arrays of thermo-optic switches to create an ADM. Polymers are probably the best materials because AWGs are temperature sensitive and polymers keep the heat localized at the switch. As a waveguide approach, thermo-optics eliminates beam alignment issues created by free-space approaches, but are susceptible to crosstalk, polarization dependence, and loss issues.

Acousto-optics uses sound waves to deflect light that passes through crystals or fused couplers. Acousto-optics will be used for wavelength-selective switches.

With holograms, electrically energized Bragg gratings are created inside a crystal. When voltage is applied, the grating deflects the light to an output port. With no voltage, the light passes straight through. Holograms compete with 3D MEMS on scalability, but are better suited to switching individual wavelengths rather than groups of wavelengths. Their nanosecond switching speed means that holograms can be used for optical packet-by-packet routing. Like LXT, holograms can support drop-and-continue applications.

2.10 Standardization Activities

Let's briefly take stock of standardization activities by the ITU, Internet Engineering Task Force (IETF), Optical Domain Service Interconnect (ODSI), and Optical Internetworking Forum (OIF), particularly with regards to the ION. They focus on the Optical UNI (O-UNI) and the Optical NNI (O-NNI).

Proponents envision intranet routers issuing commands across an O-UNI to set up lightpaths over carrier-provided networks. Some hopefully baselined their vision of architectures for commercial, service-supporting networks on the collapsed network illustrated in Figure 2-55. As enticing as the prospect is, this network vision isn't realistic and will probably never be attained in the United States, unless we decide to reestablish the Bell System or a comparable surrogate. Nonetheless, you will find commercial requirements for control that were designed against this reference model.

Figure 2-56 depicts an example of O-UNI and O-NNI in a private optical network environment.[19]

The ODSI, OIF, and other stakeholders have defined an open O-UNI that enables user devices to request optical services from the network (see Figure 2-57). This model is called an *overlay model*. The end-user device sets up static lightpaths (circuits) over the network. User protocols and network protocols are considered distinct; this hides the service provider's optical topology from the customer. ODSI completed the specification and its vali-

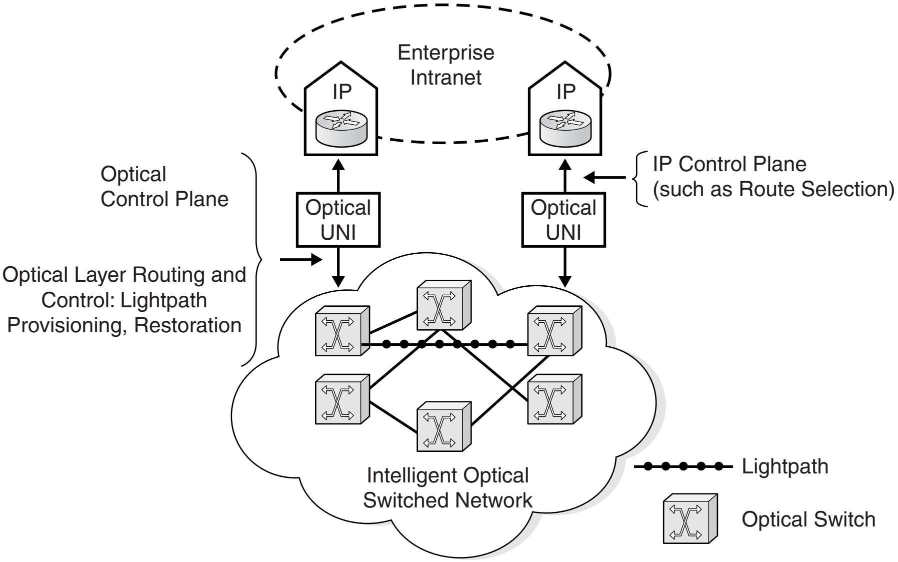


Figure 2-55 Collapsed view of U.S. carriers' metro/regional/long-haul network.

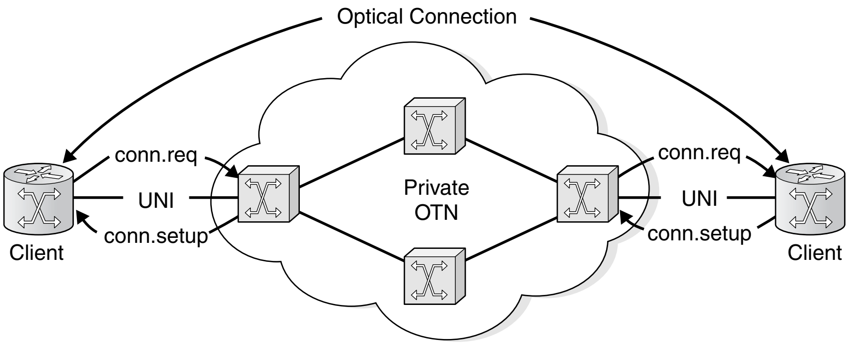
dation testing. The work was then submitted to official standards groups. In the OIF, final issues were being resolved at the end of 2001, with testing to follow. ANSI T1X1 and the ITU-T are in the requirements phase of this UNI work.

Another set of activities is underway in the IETF, which anticipates a peer-to-peer model. In this scenario, the network acts as a single collection of devices that exchange topology information, routing information, and so on (for example, using Open Shortest Path First [OSPF], Resource Reservation Protocol [RSVP], or CR-PPD). This model is supported by the GMPLS architecture and protocols. It is rather unlikely that U.S. carriers will subscribe to this kind of external control in the near future because this may well frustrate their traffic engineering, provisioning, security/access-control, and perhaps even billing efforts. In this scenario, there is a single protocol run by both user routers and the network's optical switch for circuit setup and management. This model derives directly from the MPLS model and its predecessors, Cisco's Tag Switching and Ipsilon's Flow Switching. Besides the concerns already mentioned, the model needs more complexity, less isolation, and better security.

Work to define an O-NNI is also underway. (See Figure 2-58 for NNI positioning.) Although venture capitalists may be willing to fund this work, history isn't encouraging. Neither Frame Relay intracarrier nor intercarrier

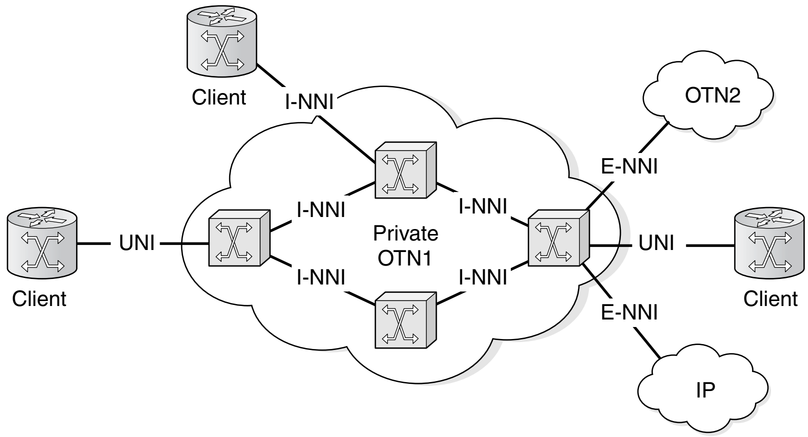
Figure 2-56 O-UNI/O-NNI in private optical networks. Source: Cisco

IP/Optical — UNI Interface



- Client-network signaling and data interface service access boundary between client(s) and OTN handles control and/or data traffic
- OTN services may be configured to signaled across UNI (analysis to PVC and SVC)

IP/Optical — NNI Interface



- Node-to-node or network-to-network adjacency, I-NNI w/in a trust domain, E-NNI between trust domains
- Characterized by equal exchange of topology data between peer entities

NNI interfaces ever really developed and took hold. The same is true of ATM intracarrier and intercarrier NNI connections.

NNIs are expected to provide connections between different carriers, subnetworks belonging to a single carrier, and the same subnetwork and switches from different vendors. This agenda is ambitious. Although it

Figure 2-57 OIF UNI 1.0 service provider and user domains.

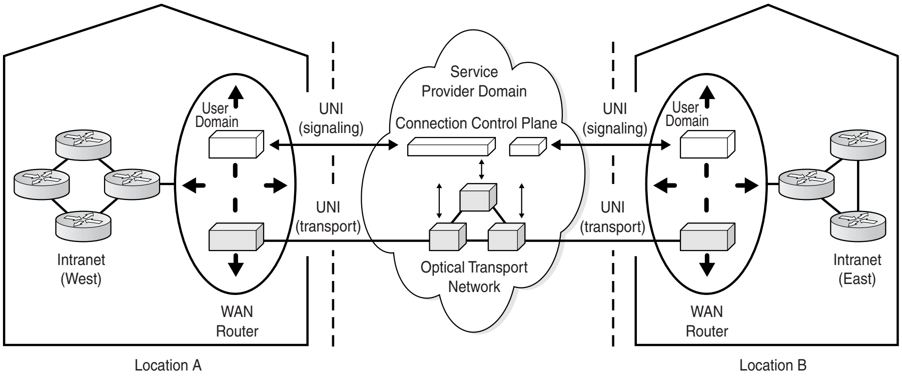
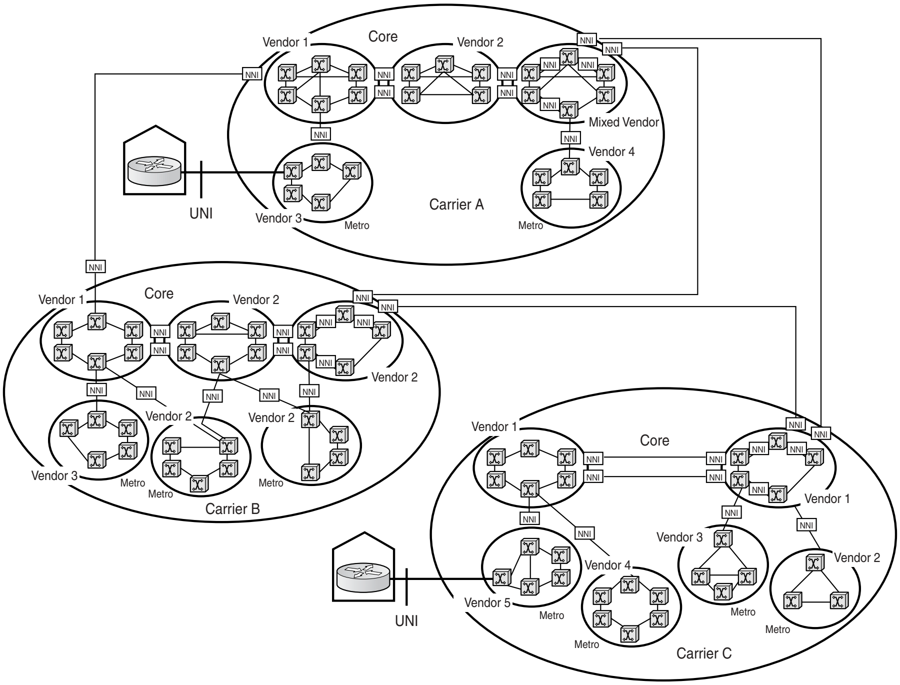


Figure 2-58 NNIs.



would seem similar to the Exterior Gateway Protocol (EGP), Border Gateway Protocol (BGP), OSPF, and MPLS, in reality those protocols generally run on the same hardware/software implementation of one vendor: Cisco IOS.

GMPLS (discussed in Chapter 8) is an extension of MPLS, which has been around for about five years and is applicable at the UNI and NNI. In this architecture, the MPLS control plane (for example, adaptations of RSVP traffic engineering [RSVP-TE] and constraint-based routing LDP [CR-LPD]) is implemented in each OXC or optical switch. The constrained-based routing mechanism and the signaling mechanism provide the control plane functionality of the OXC or optical switch. Wavelength lambda and switch ports are equated to labels; OXCs are equated to label-switched routers (LSRs) and lightpaths are equated to label-switched paths (LSPs).

End Notes

1. Although we use the term "light," this signal is in the infrared domain; hence, it is invisible to the eye.
2. Two beams are needed for duplex communication.
3. Corning introduced SMF-28 fiber, the world's most extensively deployed fiber, in 1986.
4. John George, "Optical Architectures and Fibers," IEEE EFM Meeting, Hilton Head N.C., March 12–15, 2001, Lucent Technologies.
5. K.M Feng et al., PTL (1999), and T. Day, "Widely Tunable Lasers—Tacking the Challenges of Today's Networks," www.newfocus.com.
6. Randy Bird, "Fiber Optics 101: A Primer," http://certcities.com/editorial/features/story.asp?EditorialsID=25.
7. S_2 is keyed to the S_1 character. The meaning of S_2 is different for every S_1.
8. For example, there are technical factors limiting transmission of 10 Gbps on a multimode fiber to around 82 m; new optimized multimode fibers are needed to reach (certain standards-driven goals of) 300 m.
9. Macrobending consists of loops that are much larger than the core, such as 1 inch in diameter; microbending consists of deflections of the fiber axis that are small compared to the core such as 2 nm.
10. Bellcore, *Telecommunications Transmission Engineering*, ISBN 1-878108-04-2, 1990.
11. T. Day, "Widely Tunable Lasers—Tackling the Challenges of Today's Networks," NGN 2001 Proceedings, 2001; L. A. Coldren and S. W. Corzine, "Distributed Bragg Reflector Laser," *IEEE Journal of Quantum Electronics* QE-23, 903 (1987); C. J. Chang-Hasnain, "Vertical Cavity Surface Emitting Laser," *IEEE Journal of Selected Topics on*

Quantum Electronics (June 2000): 978 ff; and P. Zorabedian and W. R. Trutna, "External Cavity Diode Laser," *Optics Letters* 13 (1988): 826 ff.

12. Composition: Indium-Gallium-Arsenide (InGaAs).

13. Composition: Germanium (Ge).

14. R. A. Barry, "Intelligent Optical Networking Year 2001: A Practical Perspective," NGN 2001 Proceedings, 2001.

15. Major new physics-related issues impact operation at the OC-768 rate.

16. Some material here is based in part on Stephen Koffler and Lori Franklin, "Optical Networking: A Focus on the Metro Opportunity," First Union Securities, New York, NY, June 2000.

17. This figure was adapted from Avanex Corporation Promotional Material, www.avanex.com.

18. This is based on Adva Optical Promotional Material, www.advaoptical.com; the rest of the section is from the same source.

19. C. Metz, "A Survey of Advanced Internet Protocols," Cisco Educational Materials, 2001.

This page intentionally left blank.

CHAPTER 3

Traditional SONET Architectures

We open this chapter with a review of the evolution of optical networks in general and the Synchronous Optical Network (SONET) in particular during the past 15 years. The rest of the chapter tackles technical aspects of SONET from the perspective of broadband metro applications and resiliency.[1]

3.1 The Evolution of Optical Networking

The deployment of fiber-optic facilities began in the early 1980s. Up to that point, the telecommunications infrastructure was built on a hierarchy that was derived from and suited to copper/coaxial transmission systems. Throughout the 1980s, these fiber-optic systems were vendor proprietary, notably based on AT&T (now Lucent) technology in the North American market. Other manufacturers introduced systems in Europe and Japan; however, none of the systems interoperated with one another. Starting in the mid-1980s, and more so in the late 1980s, the industry started to understand that standardization across vendors was desirable. In addition to being based on copper/coaxial systems, the traditional digital hierarchy suffered two additional drawbacks:

1. The speeds (up to DS3C or 90 Mbps) were not adequate for fiber.
2. The hierarchy was asynchronous (that is, it did not rely on a master clock).

SONET (North America) and Synchronous Digital Hierarchy (SDH) (the rest of the world) were developed to address these issues. SONET/SDH standards define equipment interoperability and help facilitate the translation of electrical signals to optical signals for transmission over an optical waveguide. The standards require SONET/SDH equipment to include essential multiplexing features, transmission features, and Operations, Administration, Management, and Provisioning (OAM&P) features, such as restoration, signaling, performance monitoring, and fault management. The widespread use of fiber has been made possible in part by the industry's acceptance of SONET/SDH as the standard for transmission functions (user plane) and OAM&P functions (management plane). Work on SONET/SDH standards started in the mid-1980s, but it was only in the late 1980s and early 1990s that they achieved market acceptance.

Throughout the 1990s, with the commercialization of the Internet, there was a major demand for additional bandwidth. (The growth of the Internet has been variously quoted as "100 percent a quarter"—the so-called Internet year—or more conservatively as "100 percent a year.") Using SONET/SDH standards, telecommunication companies have gradually expanded their capacity by increasing data transmission rates to the point that many carriers now routinely transport 2.5 Gbps on fiber (STM-16/OC-48), and the deployment of 10 Gbps (STM-64/OC-192) is now underway. Work on 40-Gbps STM-256/OC-768 systems is just beginning.

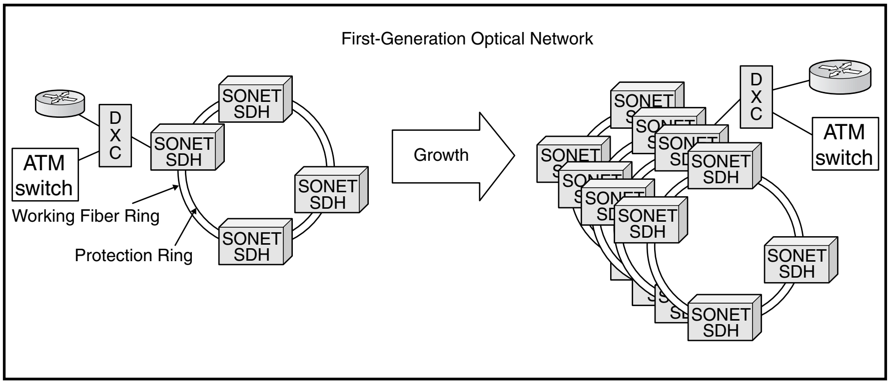
During the late 1990s, bandwidth demand was approaching the maximum capacity available in typical networks and carriers. Primarily because of the physical properties of *embedded fiber*, today there is a practical ceiling of 2.5 Gbps on many fiber networks, although there are instances where STM-64/OC-192 is also being deployed. Surprisingly, however, the time-division multiplexing (TDM) equipment installed today utilizes only a fraction of the intrinsic capacity of the fiber. Dense wavelength division multiplexing (DWDM) techniques were developed to address these issues. DWDM started to gain commercial success in the mid-1990s and achieved major deployment in the late 1990s.

Up to now, the industry has focused to a large degree on the long-haul aspect of the business. However, metro edge and metro core needs have also arisen of late and are ill served by long-haul-retrofitted solutions, which are too expensive for metro edge/metro access applications. Metro optical networks and providers need to focus in another direction—namely, access. They need to capitalize on solutions (for example, coarse WDM [CWDM]) that are more cost effective in this space, where typical distances are one to four miles. Approximately 95 percent of the office buildings in the United States, not to mention the millions of other commercial buildings, are currently not on fiber-based systems. Part of the issue relates to the cost of the in-building equipment. No one is going to make any money, much less the ill-informed venture capitalists and their investors, by deploying $80,000 to $150,000 of equipment in a building. Per-building equipment costs need to be in the $8,000 to $25,000 range for carriers to be able to deploy sustainable services.

3.1.1 First-Generation Optical Networks

As illustrated in Figure 3-1, a first-generation optical network is based on a ring architecture and uses SONET/SDH devices to transmit information

Figure 3-1 First-generation optical network.



over fiber. In this design, SONET/SDH systems interface with the fiber medium and enable a single light wave beam (a lambda or wavelength) to be carried over the network with modulation techniques that provide data rates up to 2.5 Gbps (OC-48).[2] First-generation equipment spanned the 1990 to 1995 time frame. To transmit a single wavelength, SONET add/drop multiplexers (ADMs) use a fixed source laser. The handoff to the customer is a DS3, OC-3c, OC-12c, or OC-48c, or some Layer 2 (for example, Asynchronous Transfer Mode [ATM]) or Layer 3 (for example, Packet over SONET [PoS]) interface. Typically, a digital crossconnect (DXC) at the central office (CO) or service node performs a grooming function.

This incoming user traffic is encapsulated into SONET Layer 1 frames by a Customer Premises Equipment (CPE) device (such as a SONET plug-in within a customer's router). The original Protocol Data Unit (PDU), including a Layer 2 header and Layer 2 payload (ATM header/payload), which in turn includes a Layer 3 header and Layer 3 payload (Internet Protocol [IP] header/payload), is carried from SONET node to SONET node inside the SONET envelope. Multiple SONET links may be involved in one data link layer (DLL) link; in turn, multiple DLL links are involved in an end-to-end path (for example, an IP route).

Because a standard SONET/SDH ring is limited to transmitting one wavelength (for example, 1,310 nm), there are capacity limitations in terms of what you can achieve at the fiber level. Each time the capacity of a standard ring is filled, the service provider is required to lay more fiber and add a set of SONET/SDH devices. This is costly in long-haul applications. For metro access and metro core applications, the jury is still out on the advantages of highly packed wavelengths (such as 120); however, lower-density systems (such as 8, 16, or 32) may be appropriate in the metro core space, and 2 to 4 wavelengths may be fine for metro access applications, if the cost of the devices is within specific ranges. In the metro network case, there is a traditional transmission cost versus multiplexing cost trade-off for the architecture. In long haul, the transmission cost is high; therefore, multiplexing costs are easily justified. In metro applications, the transmission cost could be bounded so the cost of electronics supporting multiplexing, particularly if this is high, may or may not always be justified.

3.1.2 Second-Generation Optical Networks

The second generation of optical networking brought the introduction of DWDM technologies to the existing SONET ring systems. In this architecture, as seen in Figure 3-2, DWDM terminals are inserted into segments of fiber rings where more than just single wavelength capacity is required. By multiplying up to 80 channels in the second generation and more in the third, ring capacity is expanded several dozen times. It should be noted immediately, however, that the cost of DWDM gear is not trivial (it could be $120,000 or more per node, with the cost being a function of the number of wavelengths supported). Therefore, although it is easily justifiable at the long-haul level, it remains a design cost issue for metro access networks and to some extent for metro core networks as well. Second-generation equipment spanned the 1995 to 2000 time frame.

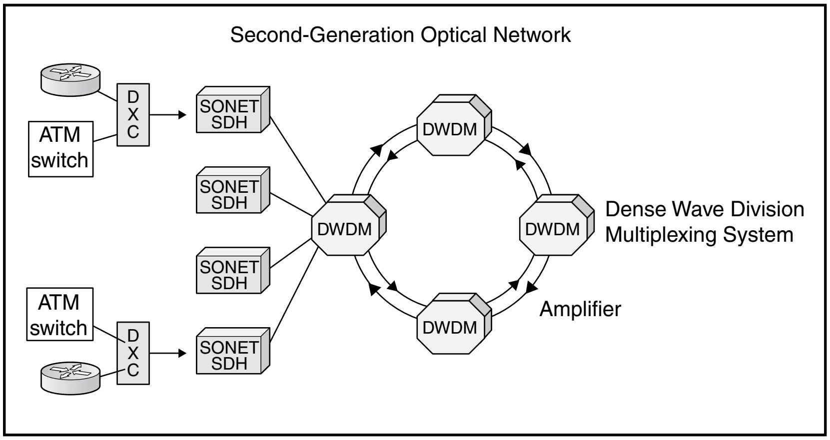


Figure 3-2 Second-generation optical network.

Typically, for each DWDM channel added, a SONET device needs to be stacked in front of the DWDM terminal to provide the wavelength source. Also, the additional capacity drives a proliferation of DXCs as more electronic traffic needs to be routed and groomed. Because of cost, the ideal application for this technology is long haul. The second generation has been described as "pushing SONET to the edge of the network." This refers to the fact that SONET is no longer connected directly to the fiber ring itself, but is connected to the DWDM terminal instead. This idea of pushing SONET to the edge of the network has further implications for the third generation of optical networking, which is described next.[3]

3.1.3 Third-Generation Optical Networks

The third generation of optical networking has evolved to deal with the cost/performance issues associated with SONET and digital crossconnect system (DCS) deployments in second-generation systems. Third-generation equipment also brought a greater level of sophistication to optical networking. Third-generation equipment spans the 2000+ time frame. At a high level, this equipment

- Incorporates the SONET, DWDM, and crossconnect functions into a fewer number of discrete, cost-reduced devices.
- Adds more intelligence to optical networking, specifically by enabling routing/switching at the wavelength level. (This is also often referred to as enabling *mesh* capability in the optical network.)

These optical networks are comprised of DWDM systems connected to intelligent optical crossconnects (OXCs). OXCs encompass the essential functions of SONET/SDH equipment, including restoration, signaling, performance monitoring, fault management, and multiplexing. Traditional SONET/SDH equipment is pushed even further to the edge of the network to manage the lower-speed interfaces and mixtures of protocols. This enables the core optical network to handle the high-speed traffic, which makes the core easy to build and manage.

It is worth noting that the traditional SONET/SDH protocol is still used in the network, only in a different way. The traditional OAM&P features are required by carriers and are performed in an efficient manner by SONET network elements (NEs). The goal of the OXC is to remove the expensive SONET/SDH equipment layer, but not the protocol itself. Another benefit of OXCs is that they support mesh architectures. A mesh network enables the

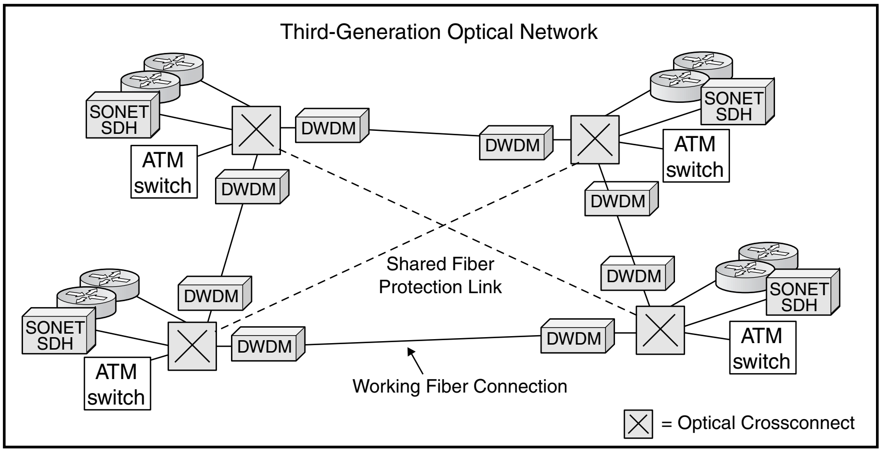


Figure 3-3 Third-generation optical network.

restoration capacity (protecting fiber) to be shared across the various spans and nodes in the network. Figure 3-3 shows that utilizing a mesh architecture lowers the amount of capacity used for protecting the network, reducing the cost of redundancy and resilience.[3]

Third-generation networks are currently being tested in the marketplace. A key feature of this network is that it has the potential to handle data rates up to 40 Gbps (OC-768) and beyond. Again, because of cost, the ideal application for this technology is still long haul.

Although equipment cost is important, it tends to be only 10 to 15 percent of the total cost incurred by a carrier and reflected in the service charges. Hence, equipment cost reduction is important, but its overall true effect has to be understood. For example, reducing a piece of equipment from $10 to $8.70 (a 13 percent decrease) only reduces the total cost of a service from $100 to $98.70, which is a mere 1.3 percent net reduction. As implied by our earlier discussion, the edge equipment cost reduction has to be in the 100 to 1,000 percent range. In our example, the cost would have to go from $10 to $5, or better yet, from $10 to $1. In the latter case, the total service cost would go from $100 to $91, or about a 10 percent reduction. Even this is not all that significant, but if the OAM&P costs are reduced from $25 to $15, then the service cost falls by 20 percent, which is where customers start to show interest in examining an alternative to the embedded service they already have. The criticality of the 100 to 1,000 percent cost reduction is made clear when carriers state that they need 67 broadband customers in a building to deliver sustainable break-even services.[4] A mere 10 to 30 percent equipment cost reduction is meaningless. This implies that disruptive technology is needed.

3.1.4 Achieving Bandwidth Capacity Goals

During the recent past, carriers have had three possible solutions to the problem of increased bandwidth demand, particularly at the long-haul level:

- Install new fiber.
- Invest in a new TDM technology to achieve higher bit rates.
- Deploy DWDM.

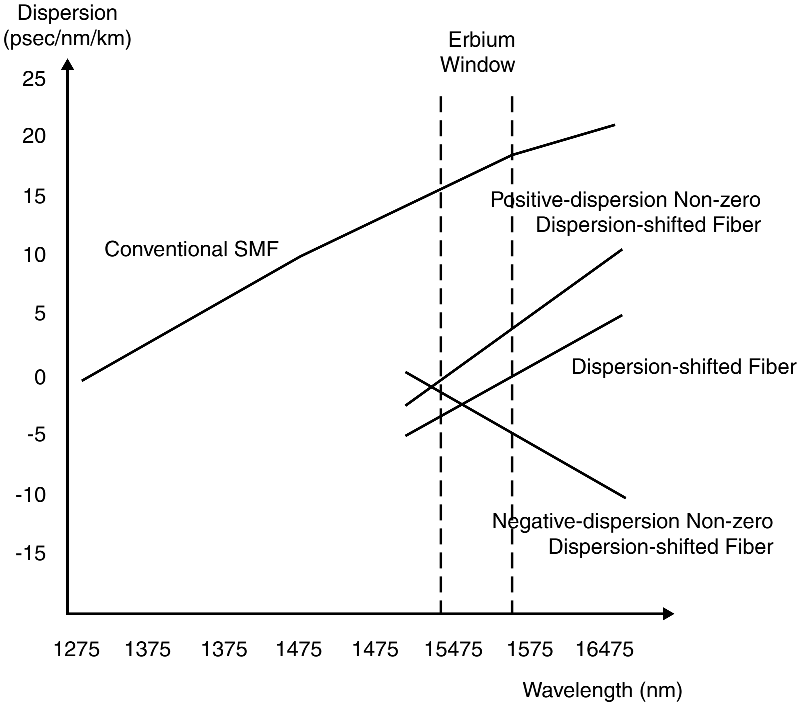
Installing New Fiber to Meet Capacity Needs Traditionally, carriers have expanded their networks by deploying new fiber and ancillary transmission equipment. For each new fiber pair deployed, the carrier could add a capacity of up to 2.5 Gbps. Unfortunately, such deployment tends to be costly. The budgetary cost to deploy the additional fiber cable, excluding the cost of associated support systems and electronics, is in the neighborhood of $50,000 to $70,000 per mile (costs are higher in downtown urban centers). Although this estimate varies from place to place, installing new fiber can be a major project. In many instances, a third party (such as a railroad or cable TV company) or even a competitor owns the right of way of the cable route or right of entry for the premises needed to house transmission equipment. For these reasons, the brute-force deployment of additional fiber is often somewhat of an impractical or last-resort solution for many carriers for both the long haul and the metro access/metro core space.

Higher-Speed TDM—Deploying STM-64/OC-192 or 10 Gbps As indicated earlier, STM-64/OC-192 is becoming an option for carriers seeking higher capacity. However, there are significant technical issues surrounding this solution, restricting its applicability. The majority of the existing fiber plant is single-mode fiber that has high dispersion in the 1,550-nm window, making STM-64/OC-192 transmission difficult to achieve reliably. In fact, dispersion has an effect that is 16 times greater with STM-64/OC-192 equipment than with STM-16/OC-48. As a result, effective STM-64/OC-192 transmission requires either some form of dispersion-compensating fiber or entire new fiber builds using nonzero dispersion-shifted fiber (NZDSF). Newer fiber can also be 50 percent more expensive than traditional single-mode fiber, impacting the $70,000 per mile just cited.

Polarization mode dispersion (PMD), which impacts the distance a light pulse can travel without signal degradation, is of particular concern for STM-64/OC-192. As transmission speeds increase, dispersion problems grow significantly, thereby reducing the distance a signal can travel; this is a major concern in long-haul applications. Specifically, PMD seems to limit the reliable reach of STM-64/OC-192 to about 50 miles on most embedded fiber (see Figure 3-4). Although there is debate within the industry over the extent of PMD problems, some key issues are known:

- PMD is particularly acute in the conventional single-mode fiber plant that comprises the vast majority of the existing fiber plant as well as in aerial fiber.
- PMD varies significantly from cable to cable and is also affected by environmental conditions, making it difficult to determine ways to offset its effect on high-bit-rate systems.
- Carriers are forced to test their constituent spans of fiber to determine their compatibility with STM-64/OC-192; in many cases, PMD will rule out its deployment altogether.

Figure 3-4
Chromatic dispersion.



A Third Approach—DWDM DWDM (discussed in greater detail in Chapter 5, "Dense Wavelength Division Multiplexing (DWDM) Systems") is a technology that enables multiple information streams to be transmitted simultaneously over a single fiber. DWDM technology utilizes a composite optical signal carrying multiple information streams, which are each transmitted on a distinct optical wavelength. The DWDM approach multiplies the baseline 2.5-Gbps system capacity by up to 80, 120, or 160 times, providing a major increase in throughput while utilizing the embedded fiber. Sixteen channel systems supporting 40 Gbps in each direction over a fiber pair are commercially available; systems are now entering the market that can support 32 channels operating at 10 Gbps each, for a total of 320 Gbps. To transmit 40 Gbps over 600 km using a traditional system would require 16 separate fiber pairs with regenerators placed every 35 km, for a total of 272 regenerators. A 16-channel DWDM system, on the other hand, uses a single fiber pair and four amplifiers positioned every 120 km, for a total of 16 regenerators to transmit over 600 km. The most common form of DWDM uses a fiber pair: one for transmission and one for reception. Systems do exist in which a single fiber is used for bidirectional traffic, but these configurations must sacrifice performance and reliability. The benefits of DWDM for increasing capacity are clear; however, this increase does not come for free. The cost can be as high as $10,000 per lambda and in some cases even more. Up to the present time, DWDM has only been cost-effective for long-haul applications.

DWDM technology has been available for over a decade, but the early applications were restricted to providing two or four widely separated wideband wavelengths. Only in the past five years has the technology evolved to the point where parallel wavelengths can be densely packed and integrated into a transmission system. By conforming to the International Telecommunications Union (ITU) channel plan (the ITU Grid), such a system ensures interoperability with other equipment.

3.2 SONET Technology

Narrowing down from the broader preliminary discussion of optical systems provided in Section 3.1, this section focuses more directly on SONET technology.

3.2.1 Overview

SONET is a North American specification for optical transport originally formulated by the industry in the mid-1980s and codified in the early 1990s into an American National Standards Institute (ANSI) standard. SONET principles were also incorporated into the SDH recommendations of the ITU, which sets standards for international telecommunications.

The basic goals of SONET include the following:

- Compatibility of equipment between vendors who manufacture to the standard (often referred to as *mid-span meet*)
- Synchronous networking
- Enhanced OAM&P
- More efficient add/drop multiplexing
- Standards-based survivable rings/meshes
- A Layer 1 infrastructure for the transport of new services, such as IP, ATM, and 10 Gigabit Ethernet (10 GbE) wide area network (WAN) PHY.

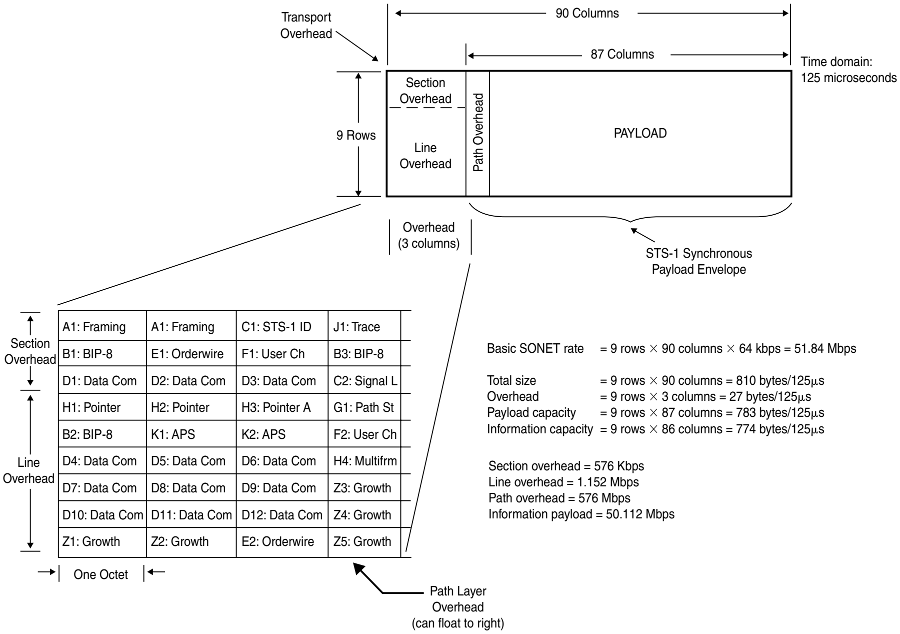
SONET defines optical carrier (OC) levels and electrically equivalent Synchronous Transport Signals (STSs) for the fiber-optic-based transmission hierarchy. The most common SONET line rates and STS-equivalent formats are shown in Figure 3-5. Multiplexing is accomplished by combining (or interleaving) multiple lower-order signals (such as 1.5 Mbps, 45 Mbps, and so on) into higher-speed circuits (such as 51 Mbps, 155 Mbps, and so on). Figure 3-6 depicts the basic frame format.

Figure 3-5
SONET optical line rates.

SONET Optical Line Rates

Optical Carrier Level	Electrical Equivalent	Line Rate Mbps
OC-1	STS-1	51.84
OC-3	STS-3	155.52
OC-12	STS-12	622.08
OC-48	STS-48	2488.32
OC-192	STS-192	9953.28

Figure 3-6
Basic frame format.



A1: Framing	A1: Framing	C1: STS-1 ID	J1: Trace
B1: BIP-8	E1: Orderwire	F1: User Ch	B3: BIP-8
D1: Data Com	D2: Data Com	D3: Data Com	C2: Signal L
H1: Pointer	H2: Pointer	H3: Pointer A	G1: Path St
B2: BIP-8	K1: APS	K2: APS	F2: User Ch
D4: Data Com	D5: Data Com	D6: Data Com	H4: Multifrm
D7: Data Com	D8: Data Com	D9: Data Com	Z3: Growth
D10: Data Com	D11: Data Com	D12: Data Com	Z4: Growth
Z1: Growth	Z2: Growth	E2: Orderwire	Z5: Growth

Basic SONET rate = 9 rows × 90 columns × 64 kbps = 51.84 Mbps

Total size = 9 rows × 90 columns = 810 bytes/125µs
Overhead = 9 rows × 3 columns = 27 bytes/125µs
Payload capacity = 9 rows × 87 columns = 783 bytes/125µs
Information capacity = 9 rows × 86 columns = 774 bytes/125µs

Section overhead = 576 Kbps
Line overhead = 1.152 Mbps
Path overhead = 576 Mbps
Information payload = 50.112 Mbps

SONET provides several capabilities that are of interest to both carriers and users:

- An electrical and optical transmission standard
- A multiplexing hierarchy
- A strong OAM&P capability
- A restoration apparatus

Figure 3-7 conveys the basic multiplexing structure of SONET. When using SONET, voice interfaces (DS1s and DS3s), high-speed data (private lines, IP, and 10 GbE), and video can be accepted by various types of service adapters. A service adapter maps the signal into the payload envelope of the Synchronous Transport Signal-1 (STS-1) or onto a virtual tributary (VT). VTs are constructs principally for the support of DS1s,[5] and this book does not focus extensively on them. New streams can be transported by a SONET system by adding new service adapters at the edge of the SONET network. All inputs are converted to an STS-1 signal. Several synchronous STS-1s are multiplexed together in either a single- or two-stage process to form an electrical STS-n signal ($n \geq 1$). STS multiplexing is performed at the byte interleave synchronous multiplexer.

Figure 3-7 Multiplexing aspect of SONET.

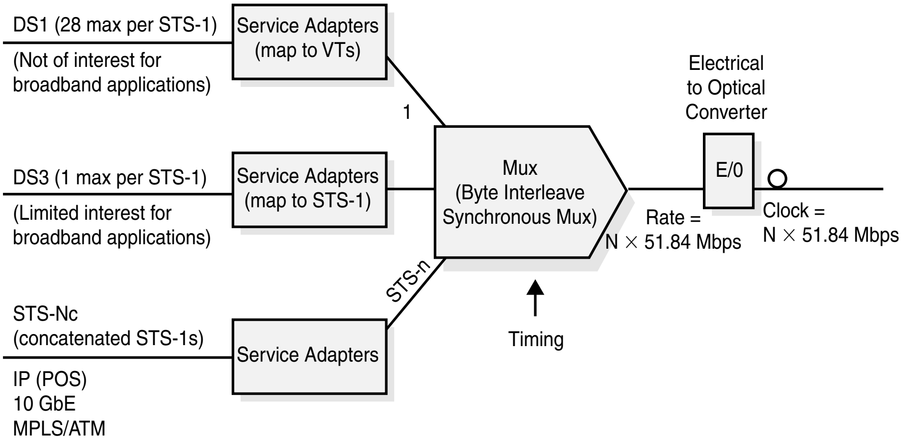


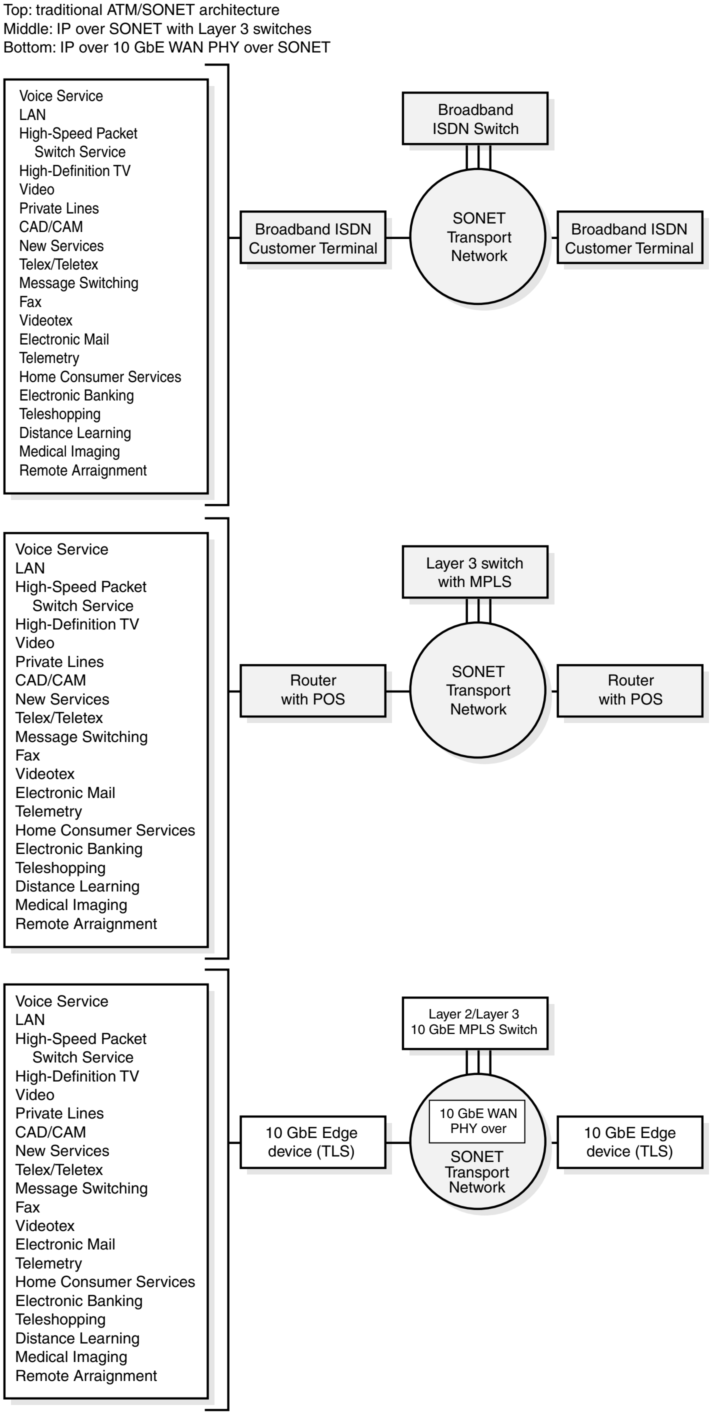
Figure 3-8 depicts the architectural evolution of private-line-like services since the mid-1990s. The first generation consisted of ATM over SONET, the second generation consisted of IP over SONET, and the third generation uses 10 GbE WAN PHY over SONET.

3.2.2 SONET Frame Format

SONET uses a basic transmission rate of STS-1, which is equivalent to 51.84 Mbps. Higher-level signals are multiples of the base rate. For example, STS-3 is three times the rate of STS-1 ($3 \times 51.84 = 155.52$ Mbps). An STS-12 rate is $12 \times 51.84 = 622.08$ Mbps. SDH is built on 155.52-Mbps modules and multiples thereof.

STS-1 Building Block The frame format of the STS-1 signal is shown in Figure 3-9, which expands upon Figure 3-6. The frame is comprised of two main areas: the transport overhead and the synchronous payload envelope (SPE). The SPE can also be partitioned into two parts: the STS path overhead and the payload. The signal is transmitted byte by byte beginning with byte 1, reading left to right from row 1 to row 9. The entire frame is transmitted in 125 microseconds. Once the payload is multiplexed into the SPE, it can be transported and switched through SONET without having to be examined and possibly demultiplexed at intermediate nodes.

Figure 3-8
Evolution of transport options.
Top: Traditional ATM/SONET architecture.
Middle: IP over SONET with Layer 3 switches.
Bottom: IP over 10 GbE WAN PHY over SONET.



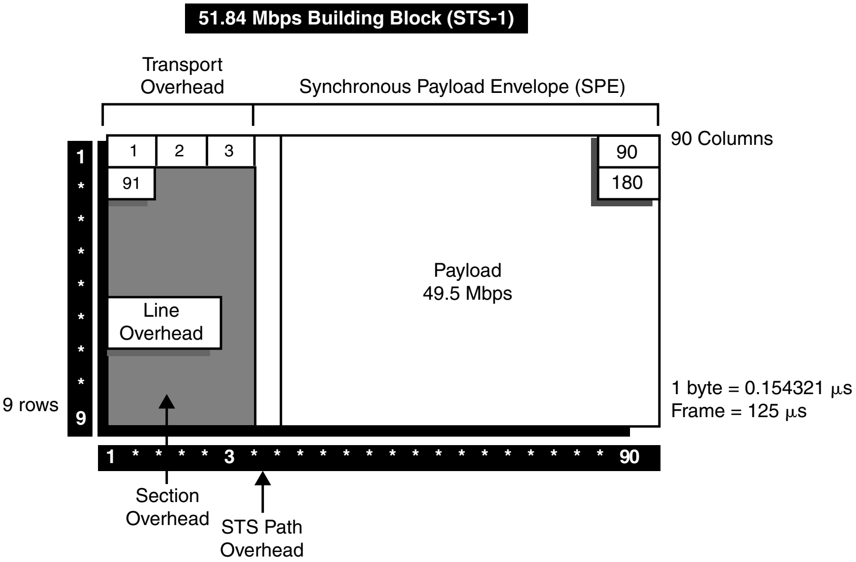


Figure 3-9
SONET STS-1 frame.

Section Overhead:
- performance monitoring (STS-n signal)
- local orderwire
- data communication channels to carry information for OAM&P
- framing

Line Overhead:
- performance monitoring of the individual STS-1s
- express orderwire
- data channels for OAM&P
- pointer to the start of the synchronous payload envelope
- protection switching information
- line alarm indication signal (AIS)
- line far end receive failure (FERF) indication

STS Path Overhead:
- performance monitoring of the STS SPE
- signal label (equipped or unequipped)
- path status
- path trace

VT Path Overhead:
- performance monitoring (virtual tributary level)
- signal label (equipped or unequipped)
- path status
- pointer (depending on VT type)

The STS-1 payload has the capacity to transport up to

- 28 DS1s
- 1 DS3
- 21 E1s (2.048-Mbps ITU-type signal) or combinations of the above

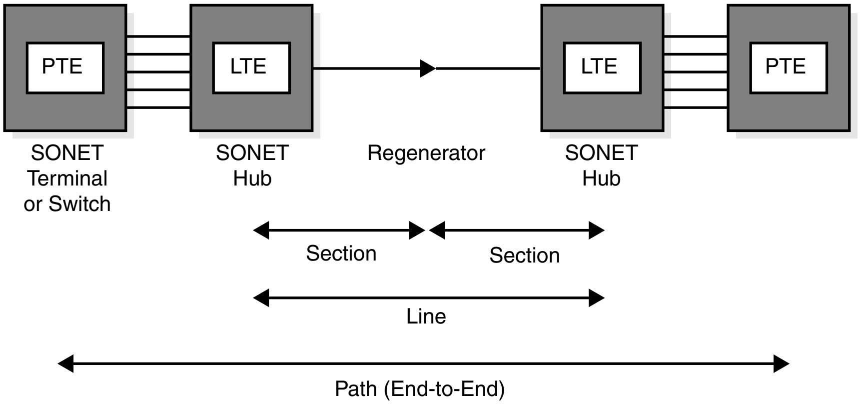
SONET carries substantial management information, expanding OAM&P capabilities compared with previous systems such as DS1 and DS3. The management information has several layers, as indicated in Figure 3-10. Path-level overhead is carried end to end; it is added to DS1 signals when they are mapped into VTs and for STS-1 payloads that travel end to end. Line overhead is used for the STS-n signal between STS-n multiplexers. Section overhead is used for communications between adjacent network elements, such as regenerators. Transport overhead is comprised of section and line overhead. The STS-1 path overhead is part of the SPE. Enough information is contained in the overhead to allow the network to operate and enable OAM&P communications between an intelligent network controller and the individual nodes.[6]

Virtual Tributaries (VTs) A rather elaborate mechanism exists in SONET for it to be able to carry (multiplex) subrate lower-speed T1, E1, and T2 channels. This mechanism is called a virtual tributary (VT). Although this has limited interest in a discussion on broadband, it is worth a quick review because it has implications on the complexity/cost of SONET equipment. The STS-1 payload may be subdivided into VTs, which are synchronous signals used to transport signals, as follows:

- **VT1.5** One DS1 (1.544 Mbps), carried in 1.728 Mbps
- **VT2** One CETP 1 (2.048 Mbps), carried in 2.304 Mbps
- **VT6** One DS2 (6.312 Mbps), carried in 6.912 Mbps

VT path overhead and VT payloads comprise the VT's SPE and are similar to the STS's payload. Within an STS-1 frame, each VT occupies a num-

Figure 3-10 SONET topological layers.



ber of columns (see Figure 3-11). The VTs are grouped in distinct groups. Within the STS-1, a number of VT groups can be mixed together to form an STS-1 payload. To accommodate different mixes of VTs in a bandwidth-efficient manner, the STS-1 SPE is divided into seven groups. A VT group may contain one VT6, three VT2s, or four VT1.5s. A VT group contains only one size of VTs, but different kinds of VT groups can be mixed into one STS-1 SPE (see Figure 3-12). Because multiplexing is synchronous, low-speed tributaries can be muxed together, but are still visible at higher rates. An individual VT containing a DS1 can be extracted without demultiplexing the entire STS-1.[7]

Pointers SONET uses pointers to compensate for frequency and phase variations. Pointers enable the transparent transport of SPEs (either STS or VT) across plesiochronous boundaries, that is, between nodes with separate network clocks having almost the same timing. The pointer is an offset value that points to the byte where the SPE begins. The use of pointers avoids the delays and loss of data associated with the use of large (125-microsecond frame) slip buffers for synchronization. Pointers provide a means of dynamically and flexibly phase-aligning STS and VT payloads, thereby permitting the ease of dropping, inserting, and crossconnecting these payloads in the network. Transmission signal wander and jitter can also be readily minimized with pointers. Figure 3-13 contains an STS-1 pointer that enables the SPE to be separated from the transport overhead. The diagram depicts a typical case of the SPE straddling over two STS-1 frames. If there are any frequency or phase variations between the STS-1 frame and its SPE, the pointer value will be increased or decreased accordingly to maintain synchronization.[6]

Figure 3-11
SONET VTs.

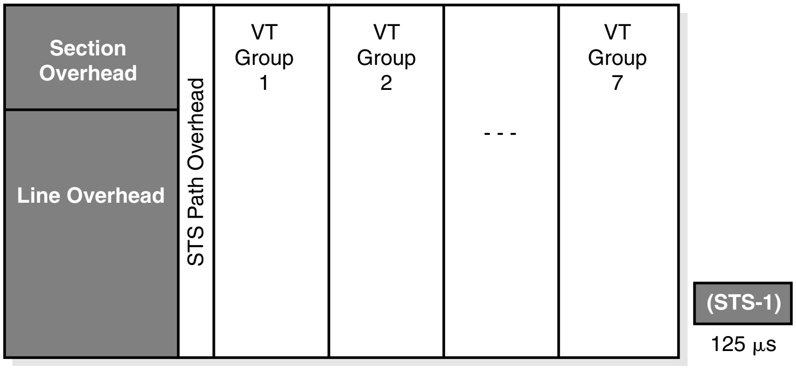


Figure 3-12
VT groups.

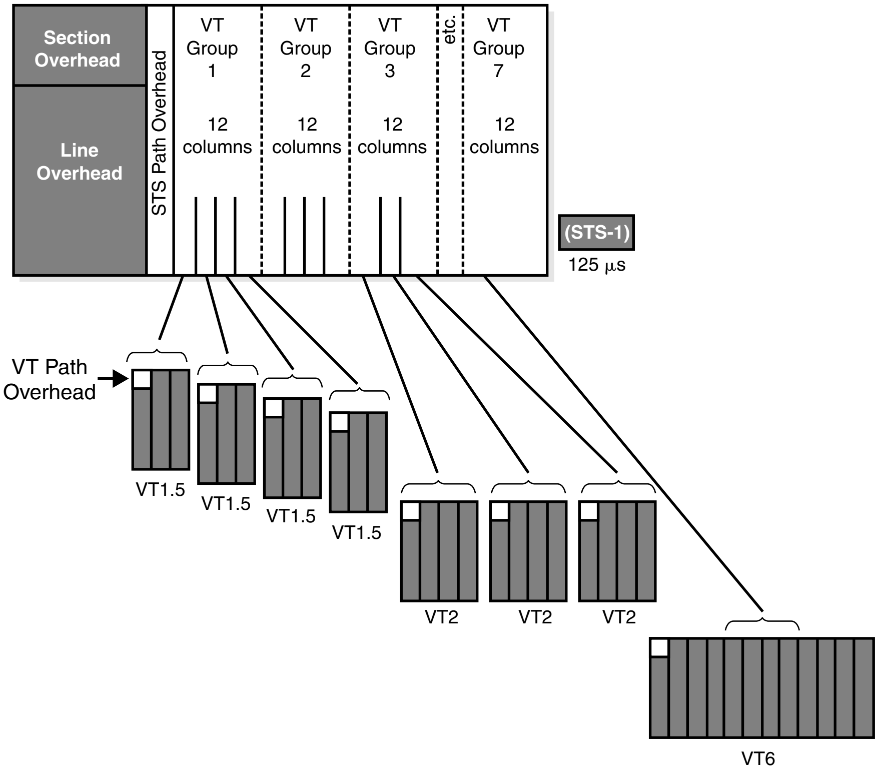
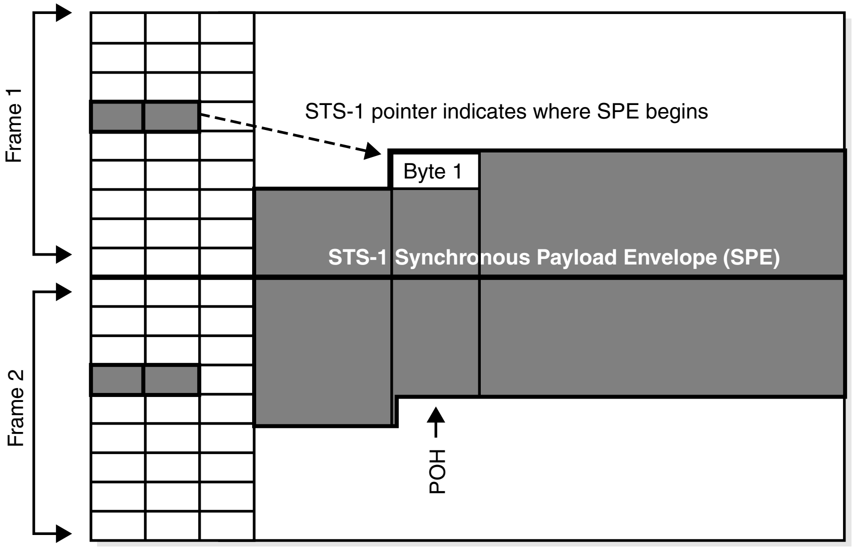


Figure 3-13
Pointer mechanism.



Concatenated Payloads For data/broadband applications, the STS-1 is short on capacity. Hence, SONET offers the solution of concatenating STS-1s to provide the necessary aggregate bandwidth. Figure 3-14 illustrates the concatenation of three STS-1s to 155.52 Mbps. Beyond STS-3, concatenation is done in multiples of STS-3c: for example, STS-12c.

3.2.3 SONET Network Elements

Although key SONET-oriented network elements are compatible at the OC-n level, vendors may include supplementary distinct features. For example, one vendor may offer an ADM with access at OC-3c only, whereas another may offer simultaneous access at OC-3c and OC-12c rates. Other vendors may offer Ethernet-based services or IP-routing functions. Figure 3-15 sketches out the key NEs.

Figure 3-14 Concatenation mechanism (OC-3c example).

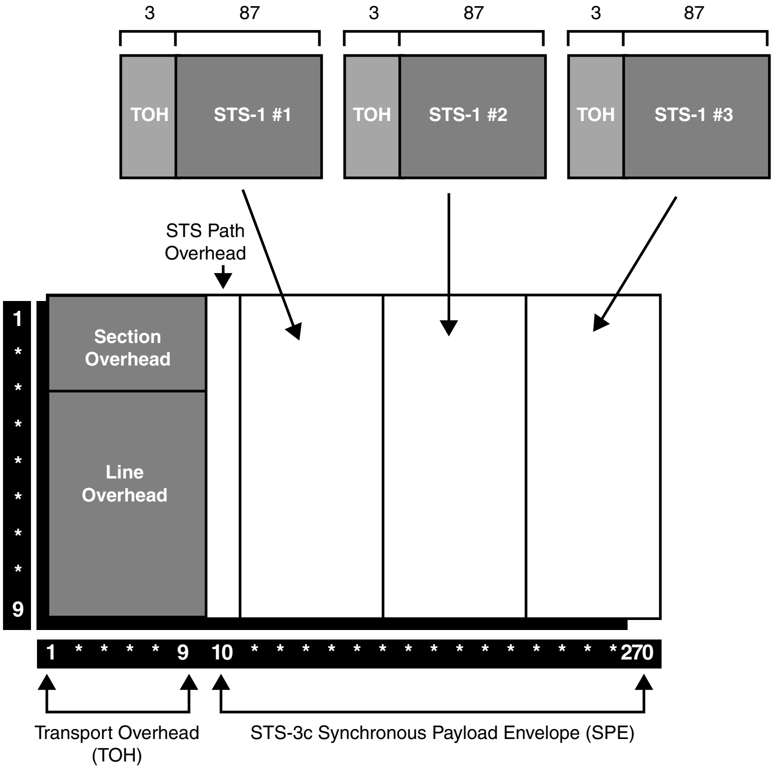
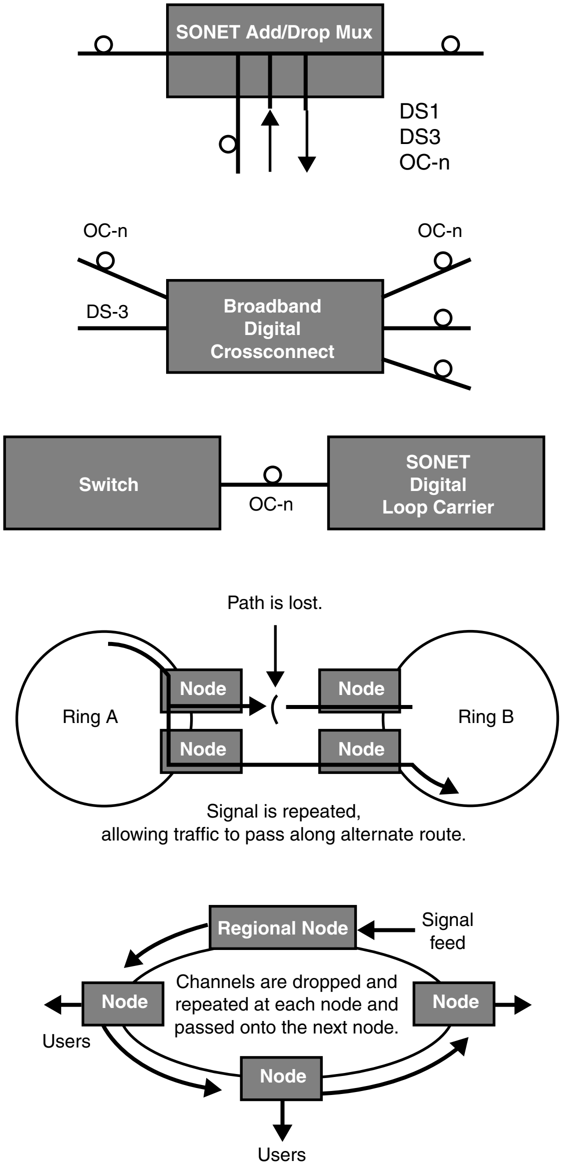


Figure 3-15
Key SONET network elements.



Add/Drop Multiplexers (ADMs)(Also Viewable as Edge Devices)

An ADM can multiplex various inputs into an OC-n signal. At an add/drop site, only those signals that need to be accessed at that site are dropped or inserted. The other traffic continues through the NE without requiring special pass-through units or other signal processing. An ADM can be deployed at a building or at an intermediate location for consolidating traffic from widely separated neighborhoods. An ADM can also be configured as a survivable ring.

Drop and Repeat SONET enables drop and repeat (also known as drop and continue). With drop and repeat, a signal terminates at one node, is duplicated (repeated), and is then sent to the next node in the ring. The next node repeats this process. This is a useful capability in telephony, cable TV, and perhaps Ethernet applications. In high-availability applications, drop and repeat provides alternate routing for traffic passing through interconnecting rings in a matched-node environment. If the connection cannot be made via one of the nodes, the signal is repeated and passed to the companion node; this node has an alternate route to the destination node.

Broadband Digital Crossconnect (DXC) A SONET crossconnect device accepts various OC rates, accesses the STS-n signals, and switches the signals at this level. (It is the synchronous equivalent of a DS3 DXC and supports hubbed network architectures.) A key difference between a crossconnect and an ADM is that a crossconnect may be used to interconnect a much larger number of STS-n signals, whereas the latter typically only drops off a single STS-n. The broadband crossconnect can be used for grooming (consolidating or segregating) STS-n signals or for broadband traffic management.

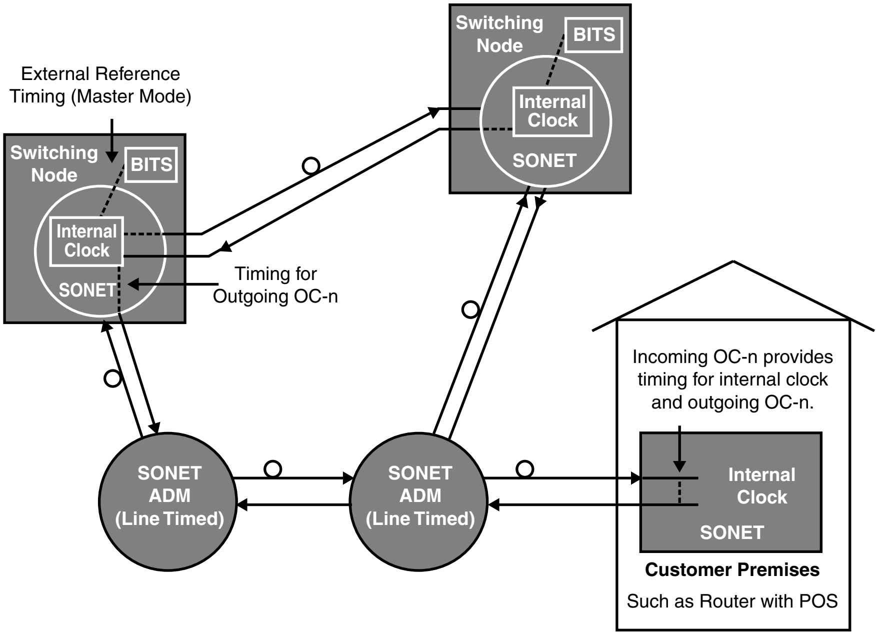
3.2.4 Synchronization

Traditionally, telco transmission systems have been asynchronous. This means that each device in the network runs on its own free-running clock.[8] Because these clocks are not synchronized, variations occur in the clock rate; this in turn impacts the signal bit rate. For example, a DS3 signal specified at 44.736 Mbps ±20 parts per million (ppm) can produce a variation of up to 1,789 bps between one incoming DS3 and another. Traditional asynchronous multiplexing uses multiple stages. Signals such as asynchronous DS1s are multiplexed. Extra bits are added to account for the variations of each individual stream (a process called *bit stuffing*) and combined with other bits (framing bits) to form a DS2 stream. Bit stuffing is used again to multiplex up to DS3. The DS1s are neither visible nor accessible within a DS3 frame. DS3s are multiplexed up to higher rates in the same manner. At the higher asynchronous rate, they cannot be accessed without demultiplexing.[6]

In a synchronous system, such as SONET/SDH, the average frequency of all clocks in the system will be the same (synchronous) or nearly the same (plesiochronous). Every clock can be traced back to a highly stable reference supply (see Figure 3-16). Hence, the STS-1 rate remains at a nominal 51.84 Mbps, enabling many synchronous STS-1 signals to be stacked together when multiplexed without any bit stuffing. Pointers accommodate differences in the reference source frequencies and phase wander, and prevent frequency differences during synchronization failures. This implies that the STS-1s are easily accessed at higher STS-n rates. Low-speed synchronous VT signals are also simple to interleave and transport at higher rates; single-step multiplexing up to STS-1 requires no bit stuffing and VTs are easily accessed.

A SONET network must be able to derive its timing from a Stratum 3 or higher clock. The internal clock of a SONET network element may derive its timing signal from a Building-Integrated Timing Supply (BITS) used by voice-switching systems and other equipment at the CO or service node. A SONET NE taking input from BITS may act as a master clock for other SONET nodes, providing timing on its outgoing OC-n signal. Other SONET nodes will operate in a slave mode called *loop timing* and have their internal clocks timed by the incoming OC-n signal.

Figure 3-16
Clock distribution.



3.2.5 SONET OAM&P

OAM&P and network management are major desiderata among carriers. SONET enjoys a high degree of OAM&P support. This has the effect of reducing the operations and support expense of a carrier; such costs are typically in the 25 percent range of the entire carrier's budget. In asynchronous networks (that is, the DS1/DS2/DS3 portion of a network, particularly when not supported by a SONET infrastructure), certain limitations exist, such as

- OAM&P bandwidth and information are limited.
- Separate operations support systems (OSSs) are used to provide maintenance functions.
- A separate data link network is used to connect OSSs to each NE.

To address these limitations, SONET

- Makes OAM&P an integral part of the transmission standard by enhancing bandwidth and information for OAM&P
- Provides more operations-level information to simplify activities such as performance monitoring
- Consists of an integral data communications network
- Makes it easier to implement centralized OAM&P

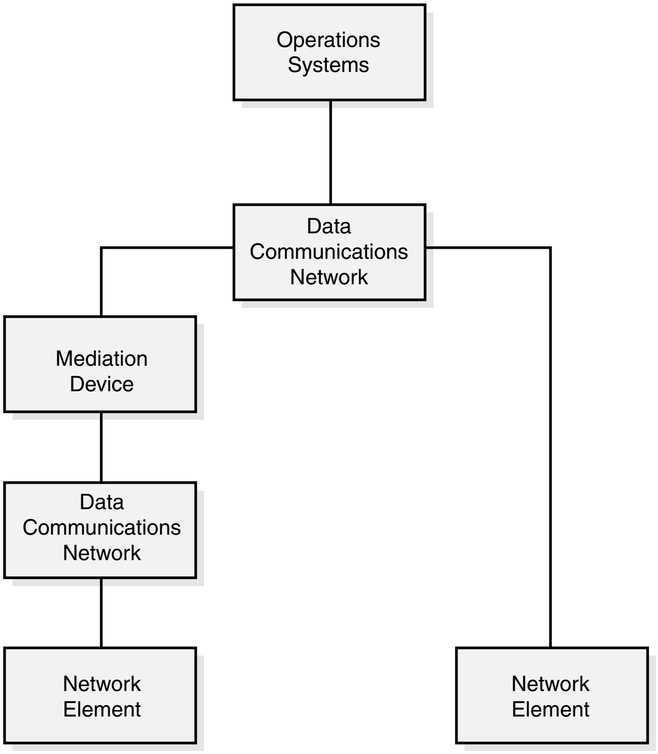
Remote provisioning and reconfiguration are vital. Because the network is continuously expanding and there are multiple vendors and types of equipment, carriers must be able to administer, monitor, provision, and control the network from a central location. Although some proponents of the new Ethernet-based metropolitan area technologies advance autodiscovery and self-provisioning, these concepts remain to be fully developed for carrier-class applications.

One of SONET's most attractive features to carriers is its capability to interface with diverse OSSs. Carriers use a constellation of OSSs to manage the many aspects of their operation, from order tracking and provisioning to engineering, from test and turnup to billing, and so on. An OSS can gather information from the network such as the alarm, status, and performance traps and statistics. An OSS can also perform remote testing or administration tasks. An OSS may also have inventory information on facilities available for provisioning and may be used to support provisioning by sending remote-control commands to enable or disable circuits. Many carriers have an extensive Operations Systems Network (OSN) that transports both the OSS-OSS, OSS-NE, and NE-NE traffic.

Standardization efforts were undertaken in the 1990s to formulate a common operation systems network architecture. The Telecommunications Management Network (TMN) Standards Committee agreed on an OAM&P architecture and formulated implementation standards for it.[9] SONET supports the TMN's architecture for OAM&P data communications. The basic architecture for OAM&P communications as defined by the ITU is shown in Figure 3-17. TMN is a framework for an OSS to communicate with the network (more accurately, its network elements), and vice versa. An OSS communicates with the network elements either directly or through mediation devices.

SONET has improved network management compared with the technology before it by providing extra bandwidth and functionality in the overhead structure to facilitate the NE-NE and NE-OSS information flow. The overhead bytes support OAM&P data channels (called data communications channels [DCCs]), enabling communications between intelligent controllers and individual network elements, as well as internode commu-

Figure 3-17 Basic TMN architecture.

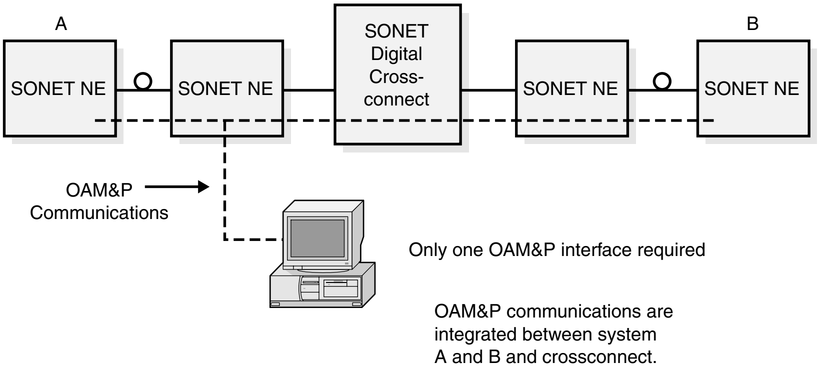


nications. This communication, although in band in the physical channel, is logically out of band; in fact, it is in a separate TDM-like channel. In contrast to this, the emerging Ethernet-based technologies either have no, or very limited, OAM&P data flows (for example, metro systems based on GbE) or are developing in-band messages commingled with the data flow, which carry only a subset of the SONET OAM&P functionality.

SONET enables circuits to be remotely enabled or disabled to carry, reroute, or remove traffic (see Figure 3-18). This capability enables the network to be dynamically reconfigured in outage situations, network restoration, traffic variations, or customer needs for time-of-day service. Other benefits include faster installation of circuits, which is important for data services; reduced need to dispatch personnel for circuit provisioning; and ease of reconfiguration, which makes the system more responsive to changes.[6] Again, although proponents of the Ethernet-based architectures advance autodiscovery and self-provisioning, these concepts remain to be fully developed at this time for carrier-class applications.

Figure 3-18 points out an important aspect of integrated OAM&P. In asynchronous networks, there are many network elements provided by different vendors. The internal overhead channels for each system are not carried from one system to another. Typically, pre-SONET FOTS[10] A and FOTS B behave as two distinct systems. Thus, separate operator interfaces are needed to perform OAM&P functions, and supportive communications are not carried through the crossconnect. In a SONET network, supportive communications are maintained from one system to the next. This capability, called *single-ended maintenance*, enables an entire network to be controlled from a single network element.[6] These concepts are also important to vendors developing new Ethernet-based metro systems. Carriers have

Figure 3-18
End-to-end OAM&P support.

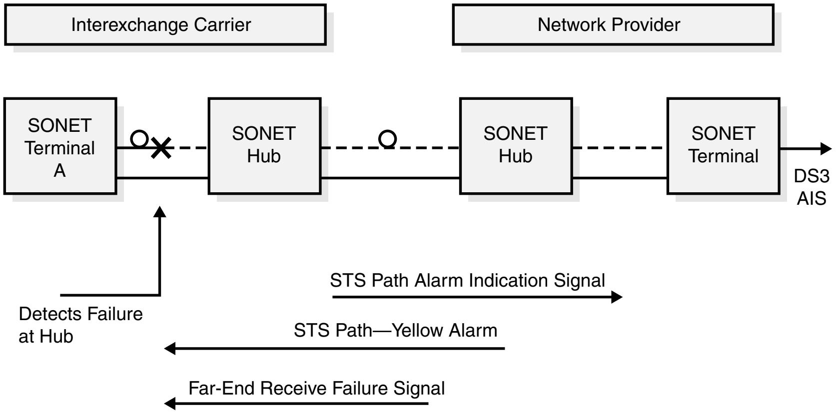


already moved away from disparate systems, particularly at the network management level. They require systems that span the entire end-to-end link, from the original edge device all the way to the termination edge device.

SONET also supports enhanced performance monitoring. As noted, SONET overhead is organized into three layers defined as section, line, and path overhead. These overhead bytes provide information for alarm surveillance and performance monitoring. A number of bytes in the SONET overhead are allocated for bit interleave parity 8 (BIP-8), a method of error monitoring. The BIP-8 is an even parity mechanism over a sequence of bits in the transmitting signal. Performance-monitoring statistics can be generated with this machinery, helping to locate degraded conditions before they become serious. BIP-8 allows for the sectionalization of alarms and single-ended maintenance. Performance information is available for all levels down to the VT path levels (via a BIP-2 capability). With SONET capabilities, the time required to restore service can be minimized. Additionally, fewer maintenance actions are required to address an impairment issue because faults can be more easily sectionalized.

An example of improved facility maintenance and surveillance appears in Figure 3-19. In case of a failure, an alarm indication signal (AIS) is sent to suppress downstream alarms. Because the terminal upstream at the left side (Terminal A) may be unaware that a failure has occurred, SONET provides a yellow AIS at the path level from the far-end terminal to Terminal A. In addition, a Far-End Receive Failure (FERF) signal at the line-layer overhead is sent from the hub detecting the failure to instruct Terminal A that a failure has occurred downstream.

Figure 3-19 Fault management in SONET.



Network survivability is imperative. Survivable rings and route diversity are crucial for the metropolitan environment. One of the key capabilities of SONET is service survivability. SONET supports both bidirectional line-switched rings (BLSRs) for internode networks and unidirectional path-switched rings (UPSRs) used in access networks. SONET-based BLSRs provide reuseable bandwidth for internode transport in meshed networks. Half the available bandwidth in a BLSR is allocated as a working route, and the other half is reserved for protection routing. For example, in an OC-48 application, the working traffic is placed in the first 24 STS-1 slots, with slots 25 through 48 serving as the protection facility. It must be noted, however, that this is an inefficient use of bandwidth.

3.2.6 Applications

The simplest implementation strategy begins by deploying SONET point-to-point systems. As the network evolves, hubs are added to the network for grooming or restoration. These may be SONET ADMs or SONET hubs. As more SONET network elements are added, the network becomes more integrated and homogenous in terms of control and OAM&P features.

Typically, operators deploy rings or meshed configuration. Occasionally, though, a single-threaded point-to-point design with pendent nodes is still employed. In a meshed network environment, the traffic is generally evenly distributed among all the nodes rather than being channeled through a small set of hubbing locations. When a traffic tributary is dropped off at one node in a BLSR, the spare capacity becomes available for reuse to pick up traffic originating at that node. If there is no new traffic, the tributary continues to another node, where it can pick up traffic originating at that location.

Figures 3-20 and 3-21 are typical metro-level networks, whereas Figure 3-22 is characteristic of a regional or national network. Urban core areas have seen the deployment of SONET-based rings and meshes, and the suburban areas follow in due time. As suggested by Figure 3-23, SONET can be deployed in the access network, especially where new fiber is being installed or existing routes are being exhausted, to provide new networking capabilities and the transport of new services. Typically, new technologies are first deployed in the interoffice portion of the network and then in the access portion.

Figure 3-20 Typical metro-level network.

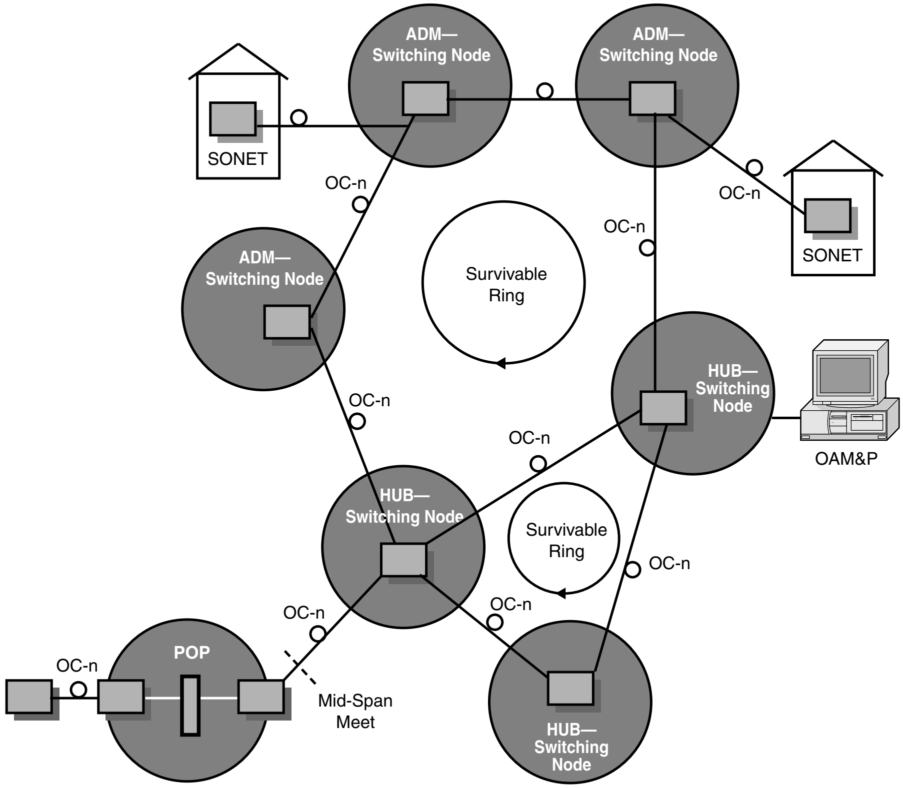


Figure 3-21 Expansion of SONET to a greater metropolitan area.

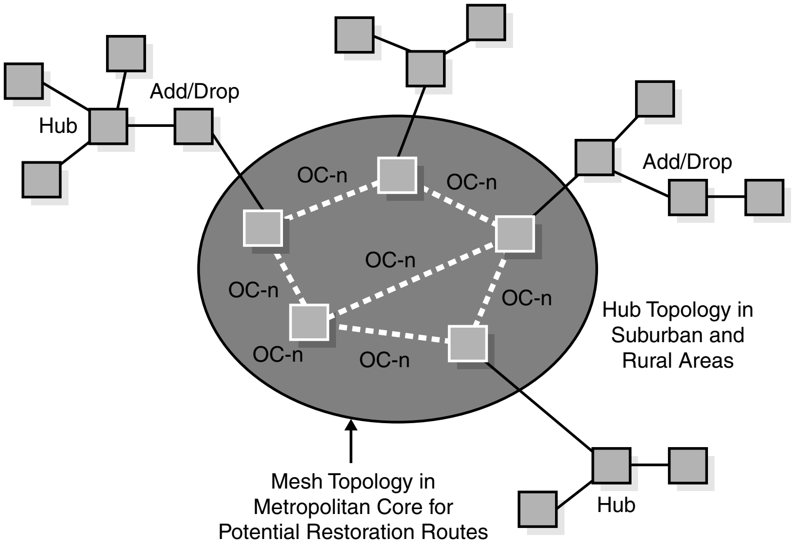
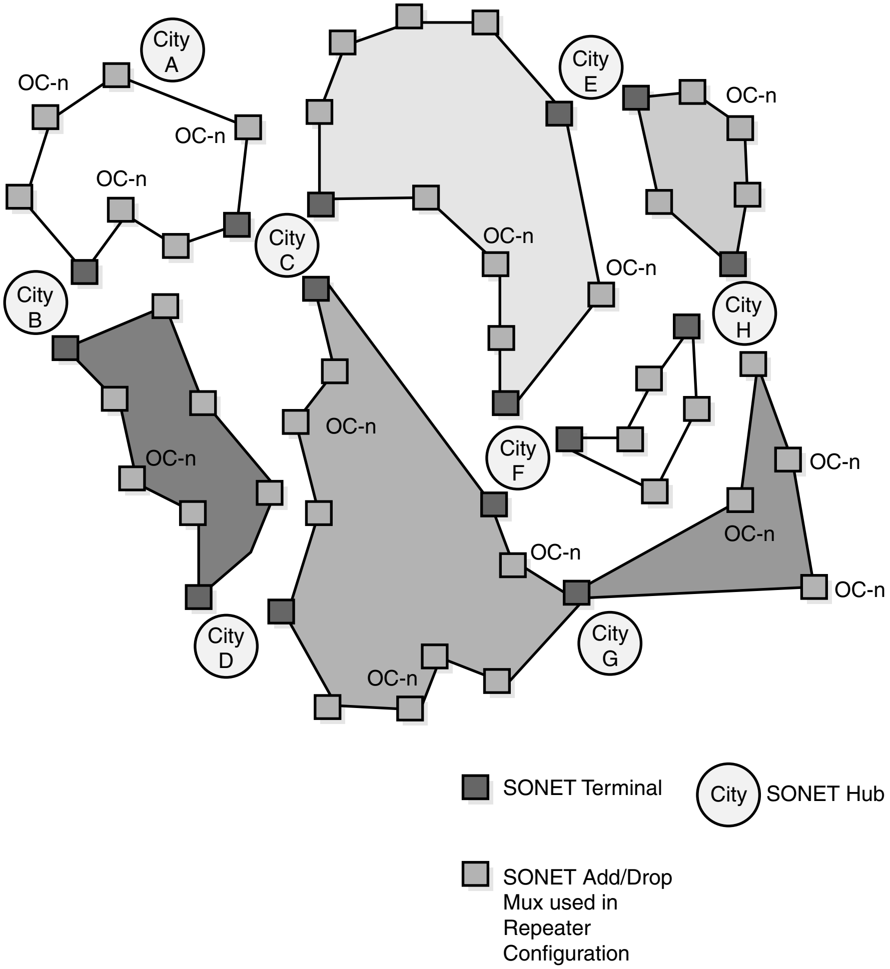


Figure 3-22
Regional or national SONET network.



3.2.7 Synchronous Digital Hierarchy (SDH)

Outside of North America, carriers adhere to SDH standards.[11] TDM systems in North America combine 24 DS0s into one 1.54-Mbps DS1 signal. European TDM multiplexes 32 E0s into one 2.048-Mbps E1 signal. (The European standard has become common in many parts of the world.) Table 3-1 compares the North American and European traditional transmission hierarchies. The original goal of the standardization effort was to efficiently accommodate both the 1.5- and 2-Mbps nonsynchronous hierarchies in a single synchronization standard. The agreement that was eventually

Figure 3-23 Point-to-point use of SONET for traditional services and ring/mesh use for new services (such as transparent LAN service [TLS]).

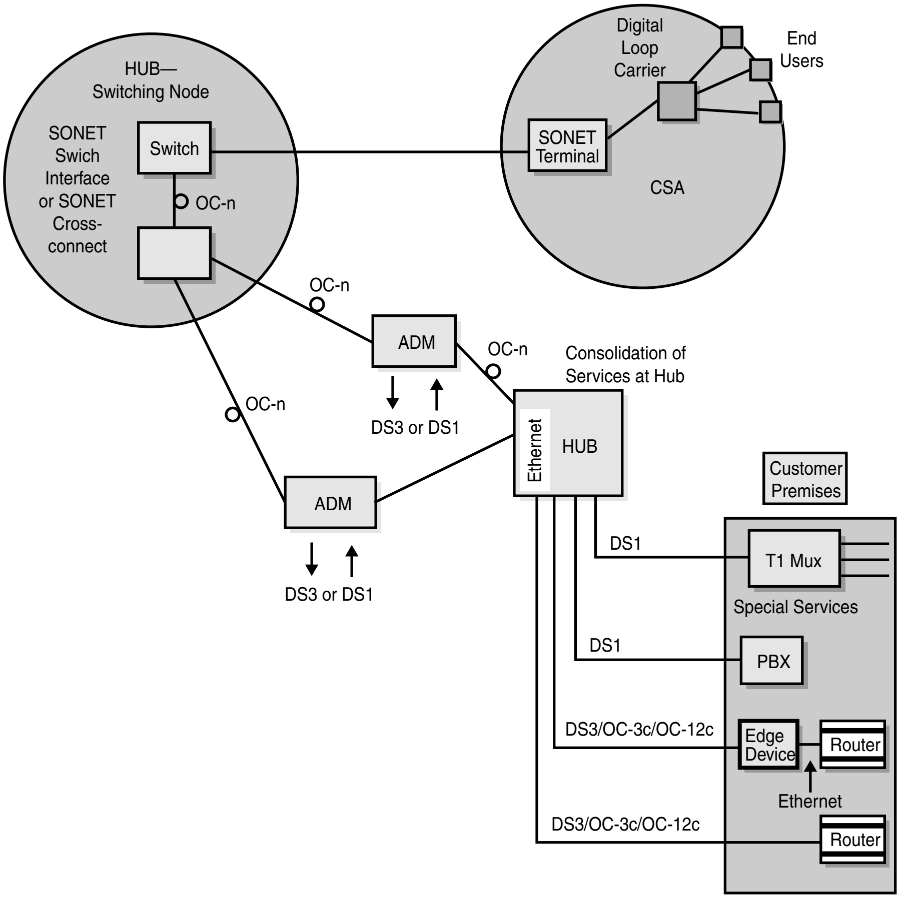


Table 3-1

Comparison of DSx/Ex Hierarchy

North American Rate			European Rate		
Signal	**Speed**	**Channels**	**Signal**	**Speed**	**Channels**
DS0	64 Kbps	1 DS0	E0	64 Kbps	1 E0
DS1	1.54 Mbps	24 DS0s	E1	2.048 Mbps	32 E0s
DS2	6.3 Mbps	96 DS0s	E2	8.45 Mbps	128 E0s
DS3	44.8 Mbps	28 DS1s	E3	34 Mbps	16 E1s
Not defined			E4	140 Mbps	64 E1s

Table 3-2

SONET/SDH Convergence

SONET	Bit Rate	SDH
STS-1/OC-1	51.84 Mbps	—
STS-3/OC-3	155.52 Mbps	STM-1
STS-12/OC-12	622.08 Mbps	STM-4
STS-24/OC-24	1,244.16 Mbps	—
STS-48/OC-48	2,488.32 Mbps	STM-16
STS-192/OC-192	9,953.28 Mbps	STM-64

reached specified a basic transmission rate of 51 Mbps for SONET and a basic rate of 155 Mbps for SDH. SONET and SDH converge at SDH's 155-Mbps base level, which is defined as Synchronous Transport Module-1 (STM-1).

The base level for SONET is STS-1 (or OC-1) and is equivalent to 51.84 Mbps. Thus, SDH's STM-1 is equivalent to SONET's STS-3 (3×51.84 Mbps$=$155.5 Mbps). Higher SDH rates of STM-4 (622 Mbps), STM-16 (2.4 Gbps), and STM-64 (9.9 Gbps) are also defined (see Table 3-2).

3.3 Standards Details

This section contains some key information from the baseline SONET standard. The basic standard of interest is ANSI T1.416-1999, "Network to Customer Installation Interfaces—Synchronous Optical NETwork (SONET) Physical Layer Specification: Common Criteria." The standard establishes common criteria for SONET interfaces at standard rates associated with the network interface (NI). Covered herein are maintenance and operation functionality at the SONET section-, line-, and path-layer specifications. Other necessary criteria for compliance with the proper interfacing of the connecting customer installation (CI) equipment are found in the other document of this series. Compliance with this standard addresses network and customer installation compatibility and is not to be construed as a constraint on the internal operations of the network or the customer installation equipment. Other documents included in the ANSI T1.416 series are listed in the following:

- ANSI T1.416.01-1999, "Telecommunications—Network to Customer Installation Interfaces—Synchronous Optical NETwork (SONET) Physical Media Dependent Specification: Multi-Mode Fiber"
- ANSI T1.416.02-1999, "Telecommunications—Network to Customer Installation Interfaces—Synchronous Optical NETwork (SONET) Physical Media Dependent Specification: Single-Mode Fiber"
- ANSI T1.416.03-1999, "Telecommunications—Network to Customer Installation Interfaces—Synchronous Optical NETwork (SONET) Physical Media Dependent Specification: Electrical"

The provisions of the ANSI T1.416 standard are intended to be consistent with applicable requirements concerning safety and environmental conditions.

Some useful definitions follow:

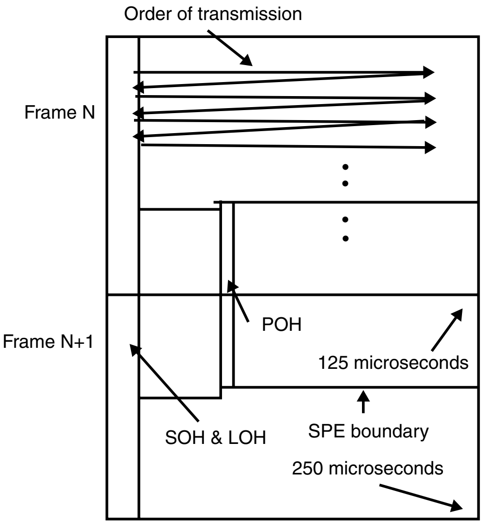
- **Scrambler** A randomizing mechanism that replaces the real data bit stream with one (at the same bit rate) that can be conceived as the process of exclusive ORing the data stream with delayed copies of itself (or self-synchronous scramblers) or exclusive ORing that data stream with a bit stream created by the exclusive ORing of an initial pattern with delayed copies of itself (for frame synchronous scramblers). The process is often represented by generation polynomials. The intent is to eliminate long strings of ones and zeros to improve the effectiveness of the clock synchronization. The received signal is descrambled by the inverse process.
- **VT1.5 interface line-terminating element (VT1.5LTE)** Network elements that originate and terminate VT1.5 interface signals. VT1.5 LTEs can originate or terminate the line overhead of the VT1.5 interface signal.
- **VT1.5 transport format** The VT1.5 transport format has a 2.048-Mbps format. This format has 32 bytes per 125 μsec frame, allocating 26 bytes for single VT1.5 SPE, 1 byte for a VT1.5 pointer, and 5 bytes per frame of line overhead. This format multiplexes the 1.728-Mbps bandwidth for a single VT1.5 payload and pointer as well as 320 Kbps of line overhead into a 2.048-Mbps transport format.
- **VT1.5 interface** The SONET VT1.5 transmission interface that transports one VT1.5 SPE and pointer (27 bytes per frame) along with a line overhead (5 bytes per frame). VT1.5 interfaces may be defined by several electrical and optical physical transport technologies.

3.3.1 Common Criteria

The order of transmission of bits and bytes is specified in ANSI T1.105. For all diagrams of frame structures shown in this standard, bytes will be transmitted in the following order for each frame:

1. The uppermost row is transmitted first, byte by byte, from left to right.
2. The second row is transmitted from left to right.
3. The succeeding rows are transmitted in sequential order, from top to bottom, as depicted in Figure 3-24.

Figure 3-24
SONET frame.



Section overhead (SOH)
A1: Frame alignment
A2: Frame alignment
B1: Section BIP-8
D1-D3: Section data communication channel
E1: Orderwire
F1: Section user channel
J0: Section trace
Z0: Section growth

Line overhead (LOH)
B2: Line BIP-8
D4-D12: Line data communication channel
E2: Orderwire
H1 and H2: Pointer value
H3: Pointer action byte
K1 and K2: APS channel
M0: STS-1 line REI
M1: STS-N line REI
S1: Synchronization messaging
Z1: Growth
Z2: Growth

STS path overhead (POH)
B3: Path BIP-8
C2: STS path signal label
F2: Path user channel
G1: Path status
H4: Multiframe indicator
J1: STS path trace
N1: Tandem connection maintenance/path data channel
Z3 and Z4: Growth

VT path overhead
J2: VT path trace
V5: VT path
Z6 and Z7: VT path growth

VT line overhead
M1: Line quality
S1: Synchronization message

3.3.2 SONET Overhead

Detailed definitions for the following SONET overhead can be found in ANSI T1.105. The following paragraphs provide an abbreviated listing of the available overhead bytes. Not all bytes will be utilized on customer-to-network applications. The SONET overhead is active across the NI, as shown in Figure 3-25.

The section overhead is shown in Figure 3-26, the STS path overhead is shown in Figure 3-27, the line overhead is shown in Figure 3-28, and the VT overhead is shown in Figure 3-29.

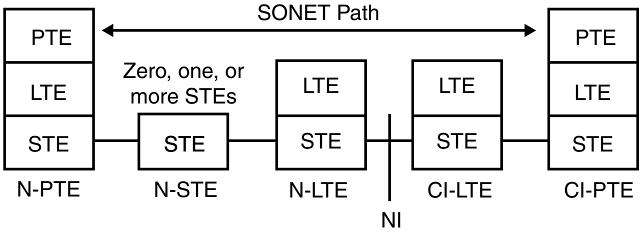
3.3.3 Synchronization

In normal synchronous operation, the timing of the SONET frame at all interfaces points on the network heading in the direction of the customer installation needs to be traceable to a Primary Reference Source (PRS) as specified in ANSI T1.101. Also, in normal synchronous operation, the timing of the SONET frame at all NIs points heading in the direction away from the customer installation (CI)[12] needs to be traceable to the PRS.

During normal operation, the customer installation transmits a signal that has a bit rate accuracy equal to that of the received signal by either of the following methods:

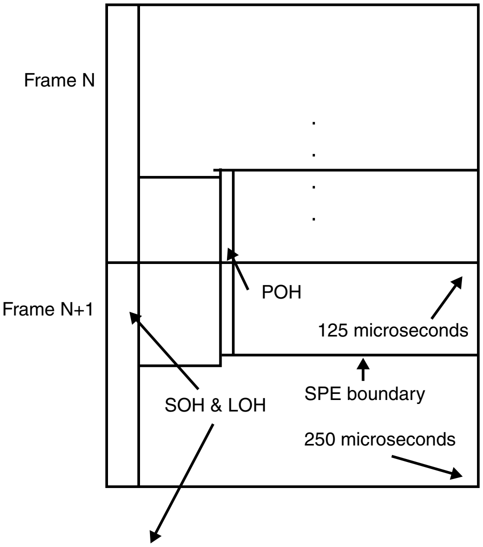
- Locking the frequency of its transmitted signal clock to the long-term average of the incoming signal. (This is often referred to as *loop timing*.)
- Providing equal signal bit rate accuracy from another PRS.

Figure 3-25 Example of SONET equipment associated with the NI.



PTE = Path Terminator Equipment
LTE = Line Terminating Equipment
STE = Section Terminating Equipment

Figure 3-26
Section overhead.

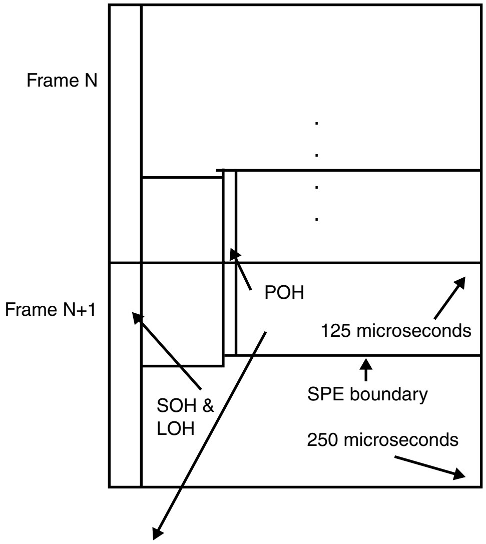


Overhead byte	Function	Usage	Coding
Section overhead			
A1 framing	Frame alignment	R	11110110
A2 framing	Frame alignment	R	00101000
B1 section BIP-8	Section error monitoring	R	BIP-8
D1-D3 section data communication channel (Data Com)	Section data communications channel	O	
E1 orderwire	Orderwire	O	
F1 section user channel	Section user channel	O	The F1 byte is defined only for STS-1 number 1 of an STS-N signal.
J0 section trace	Section trace	O	Defined only for STS-1 number one in an STS-N signal. When the Section Trace function is not supported or if no value has been programmed, then 01 Hex is transmitted.
Z0 section growth	Section growth	O	One byte is defined in each STS-1 for future growth except for STS-1 number 1 (which is defined as J0).

R = required; A = application-specific function; O = optional

When in a maintenance state (that is, a synchronization failure condition), the timing of the signals from the network may lack traceability to a PRS. In that case, the signals at all interface points heading in the direction from the network to the CI need to be within ±20 ppm of the nominal rate specified for each interface. Also, when in a maintenance state, the timing

Figure 3-27
Path overhead.

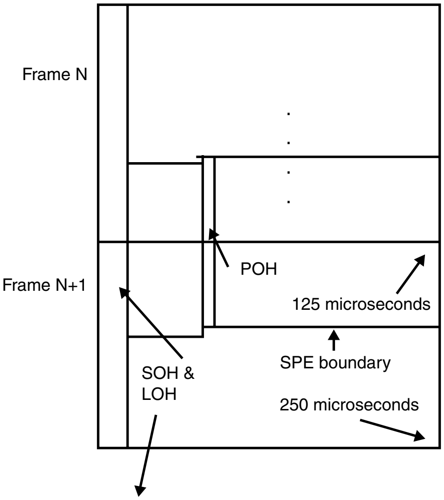


Overhead byte	Function	Usage	Coding
STS path overhead			
B3 path BIP-8	Path error monitoring	R	BIP-8.
C2 STS path signal label	STS path signal label	R	Use as defined in ANSI T1.105.
F2 path user channel	Path user channel	A	Required for use by PTE.
G1 path status	Path status	R	One byte is allocated to convey back to an originating STS PTE the path terminating status and performance.
H4 multiframe indicator	Multiframe indicator	A	Only required for VT-structured payloads.
J1 STS path trace	STS path trace	R	The content of the message is not constrained by this standard, since it is assumed to be user programmable at both the transmit and receive ends.
N1 tandem connection maintenance/path data channel	Tandem connection maintenance/ path data channel	A	
Z3-Z4 growth	Growth	O	Two bytes are allocated for future as-yet-undefined purposes.

R = required; A = application-specific function; O = optional

of the signals heading from the CI to the network may lack traceability to a PRS. In that case, the signals at all interface points heading in the direction from the CI to the network must be within ±20 ppm of the nominal rate specified for each interface.

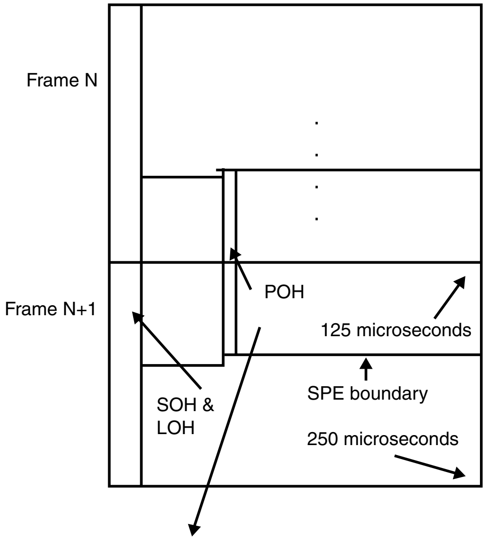
Figure 3-28
Line overhead.



Overhead byte	Function	Usage	Coding
Line overhead			
B2 line BIP-8	Line error monitoring	R	BIP-8 for STS-1 on each STS-1 of STS-N. BIP-8 for STS-Nc (BIP-24 for STS-3c).
D4–D12 line data communication channel (Data Com)	Line data communications channel	O	
E2 orderwire	Orderwire	O	
H1–H2 pointer	Pointer	R	These bytes will be provided in all STS-1 signals within an STS-N signal. Concatenation is handled as defined in ANSI T1.105.
H3 pointer action byte	Pointer action	R	These bytes will be provided in all STS-1 signals within an STS-N signal.
K1, K2 APS channel	Automatic protection switch	A	If APS is used, will operate as defined in ANSI T1.105.01.
M0 STS-1 line REI	STS-1 line REI	R	STS-1 line only.
M1 STS-N line REI	STS-N line REI	R	In a SONET signal at rates at or above STS-3, one byte, the M1 byte, is allocated for a line REI function. The M1 byte is located in the third STS-1 in order of appearance in the byte interleaved STS-N frame.
S1 synchronization messaging	Synchronization messaging	R	Defined only for STS-1 number one in an STS-N signal.
Z1 growth	Line growth	O	One byte is defined in each STS-1 for future growth except for STS-1 number 1.
Z2 growth	Line growth	O	In SONET signals at rates above STS-1 and below STS-192, one Z2 byte is defined in each STS-1 except the third STS-1 for future-growth interleaved STS-N frames.

R = required; A = application-specific function; O = optional

Figure 3-29
VT overhead.



Overhead byte	Function	Usage	Coding
VT path overhead			
J2 VT path trace	VT path trace	O	For further study.
V5	Path error monitoring, path signal label, and remote defect monitoring	R	Use as defined in ANSI T1.105.
Z6 VT path growth	Growth	O	Allocated for future as-yet-undefined purposes.
Z7 VT path growth	Growth and remote defect monitoring	A	Use as defined in ANSI T1.105.
VT line overhead			
M1 line quality	Line quality	R	One byte is allocated every superframe (16 kbit/s) for the line quality functions. Line quality functions include the VT1.5 line BIP-2, two-bit line REI, one-bit RDI-LV, one framing check bit, and two reserved bits.
S1 synchronization message	Synchronization message	R	One byte per superframe (16 kbps) is allocated for a synchronization message function. Bits 5-8 of this byte will be allocated to carry the same four-bit synchronization message codes as defined for other SONET transport interfaces. Bit 2 is a framing check bit. Bits 1, 3, and 4 of this byte are reserved.

R = required; O = optional

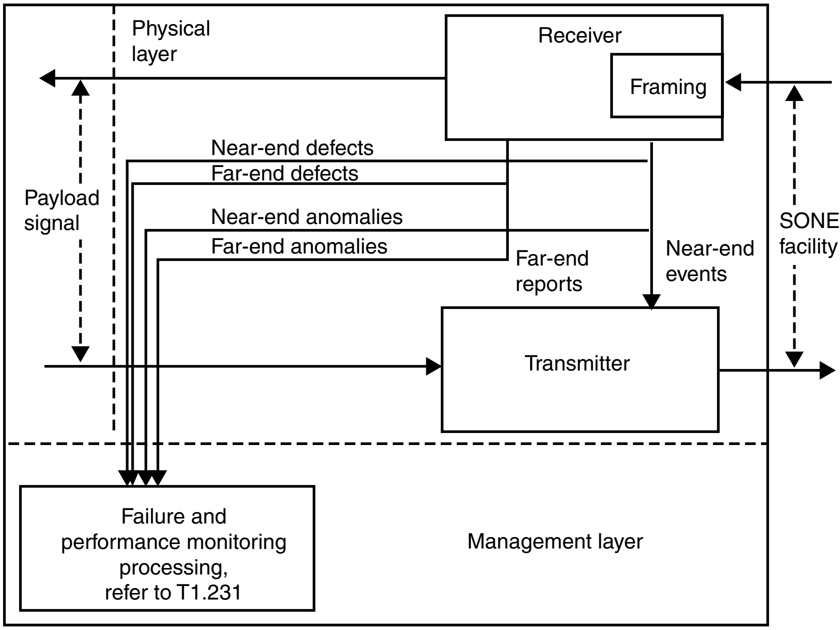
Although the previous example represents minimum requirements on synchronization for SONET NIs, it does not address all of the complexities associated with network synchronization and interfaces. For further information, refer to ANSI T1.105.09. Some of the aspects of SONET synchronization and timing that are dealt with in ANSI T1.105.09 are as follows:

- The effects of timing anomalies on SONET payload pointers and their respective payloads
- Synchronization performance expected during the fault transition from a fully synchronous operation to a worst-case, free-run clocking condition (for example, what clock holdover characteristics are specified for SONET network elements)
- Stratum compatibility criteria

3.3.4 Maintenance Physical-layer maintenance is accomplished by monitoring the received signals (see Figure 3-30). This figure illustrates the way monitored primitives are used for failure and performance tracking. Near-end events and far-end reports in terms of anomalies, defects, and

Figure 3-30 Physical-layer maintenance functions.

ANSI T1.416-1999



failures are discussed in the following section. Events and reports are used for two purposes:

- All near-end events and far-end reports will be processed for performance-monitoring primitives and failures, as described in ANSI T1.231.
- Near-end events will be used to make up far-end reports for the transmitted signal.

SONET facilities terminate at three hierarchical levels:

- Section
- Line
- Path

Maintenance capabilities are performed by SONET network elements and are made possible by maintenance tools built into the overhead fields of the SONET framing structure. Performance monitoring and failure processing for the physical media, section, line, and path are specified in ANSI T1.231. Near-end events and failures used to make up far-end reports are also specified in ANSI T1.231. Near-end events and failures used to make up far-end reports for the transmitted signal are specified in ANSI T1.105. The near-end events and far-end reports available in SONET are summarized in Figure 3-31.

Physical Medium

STS-1 Rate and Above A loss of signal (LoS) defect occurs upon the detection of no transitions on the incoming signal (before descambling) for time T, where $2.3 \leq T \leq 100$ μs. The LoS defect is terminated after a time period equal to the greater of 125 μs or $2.5 \times T'$ containing no transition-free intervals of length T', where $2.3 \leq T' \leq 100$ μs.

VT1.5 Rate—Electrical Interface A Loss of Signal/Line VT Interface (LoS-LV) defect occurs when no transitions on the incoming signal for a time period equal to 175 ± 75 contiguous pulse positions are detected.

A LoS-LV defect is terminated upon detecting at least 20 percent of the transitions on the incoming signal over a time period equal to 175 ± 75 contiguous positions.

Figure 3-31 Near-end events and far-end reports.

Primitives for failure and performance monitoring							
	Physical media	Section		Line		Path (facility)	
	Defect	Anomaly	Defect	Anomaly	Defect	Anomaly	Defect
Near-end	LoS	BIP-N (S)	SEF/LoF	BIP-N (L) BIP-2 (LV)	AIS-L SEF-LV LoF-LV	BIP-N (P)	LoP-P AIS-P TIM-P UNEQ-P
Far-end				REI-L REI-LV	RDI-L RDI-LV	REI-P	RDI-P OR ERDI-P

LoS = Loss of Signal
BIP-N = Bit-Interleaved Parity-N
SEF = Severely Errored Frame
REI-L = Remote Error Indication-Line
REI-LV = Remote Error Indication-Line, VT Interface
AIS = Alarm Indication Signal
SEF-LV = Severely Errored Frame/Line, VT Interface
LoS-LV = Loss of Signal/Line VT Interface
RDI-L = Remote Defect Indication-Line
RDI-LV = Remote Defect Indication-Line, VT Interface
TIM-P = Trace Identifier Mismatch-Path
UNEQ-P = Unequipped-Path
ERDI-P = Enhanced RDI-Path

Section The following primitives apply to the STS-N section level:

- **Section BIP-8 error** A BIP-8 error event is an anomaly that occurs when the locally computed BIP-8 does not agree with the BIP-8 sent by the signal source.

  A BIP code represents the even parity of a given signal for each bit position. Section BIP-8 is computed over all bits of the previous STS-N(c) frame after scrambling. The computed BIP-8 is placed in the first B1 byte of the next frame before scrambling.

- **Loss of frame (LoF)** An LoF defect occurs when a severely errored frame (SEF) persists for a period of 3 ms. The LoF defect is terminated when no SEF defects are detected for a period T, when 1 ms$\leq T \leq$3 ms.

  An SEF defect is the occurrence of four or five contiguous errored frame alignment words. A frame alignment word occupies the A1 and A2 bytes of an STS frame. An SEF defect is terminated when two contiguous error-free frame words are detected.

Line Two rates are considered here:

STS-1 Rate and Above

- **Line BIP-N*8** A line BIP-N*8 error event is an anomaly that occurs when the locally computed BIP-N*8 does not agree with the BIP-N*8 sent by the signal source.

  Line BIP-N*8 is computed over all bits of the previous STS-N(c) except for the first three rows of the section overhead and is placed in the B2 bytes of the current S-N(c) frame before scrambling.
- **Remote error indication-line (REI-L)** The occurrence of an REI-L>0 is an anomaly. REI-L is a count of the difference between the computed line BIP-N*8 and the received B2 bytes. The count is placed in the M0 byte (STS-1 line signal) or M1 byte (STS-N line signals where $N>1$) in the opposite direction of transmission.
- **Alarm indication signal-line (AIS-L)** An AIS-L defect is the occurrence of line AIS in five contiguous frames. An AIS-L defect is terminated when no line AIS is detected in five contiguous frames.

  Line AIS is an STS-N(c) signal containing a valid section overhead with bits 6, 7, and 8 of the K2 byte set to 111, and a scrambled all-1s pattern for the remainder of the signal.
- **Remote defect indication-line (RDI-L)** An RDI-L defect is the occurrence of an RDI-L signal in x contiguous frames, where $x = 5$ or 10. An RDI-L defect is terminated when no RDI-L signal is detected in x contiguous frames.

  An RDI-L signal is 110 in bits 6, 7, and 8 of the K2 byte in STS-1 number 1. It is set by the far-end source when it detects an AIS-L defect.

VT1.5 Rate

- **Line BIP-2** A BIP-2 error event is an anomaly that occurs when the locally computed BIP-2 does not agree with the BIP-2 sent by the signal source.

  Line BIP-2 is computed over all bits of the previous frame and is placed in bits 5 and 6 of the VT1.5 line quality byte (M1).
- **REI-LV** The occurrence of an REI-LV>0 is an anomaly. REI-LV is a count of the difference between the computed line BIP-2 and the received BIP-2 in the line quality byte (M1). The count is placed in bits

7 and 8 of the VT1.5 line quality (M1) in the opposite direction of transmission.

- **SEF-LV** Under study.
- **LoF-LV** Under study.
- **RDI-LV** The RDI-LV defect occurs when the presence of the RDI-LV signal in 10 contiguous frames is detected. The RDI-LV defect is terminated when no RDI-LV signal is detected in 10 contiguous frames.

  The RDI-LV signal is the occurrence of a 1 in bit 4 of the line quality byte (M1).

  Figure 3-32 provides a view of the VT1.5 superframe.

Path The following primitives apply to the path level:

- **Path BIP-8** A BIP-8 error event is an anomaly that occurs when a locally computed BIP-8 does not agree with the BIP-8 sent by the signal source.

  Path BIP-8 is calculated over all bits of the path overhead and payload before scrambling. The computed BIP-8 is placed in the B3 byte of the following frame before scrambling.
- **REI-P** The occurrence of an REI-P>0 is an anomaly. REI-P is the count of the difference between path BIP-8 received in B3 after unscrambling and the path BIP-8 calculated over path overhead and

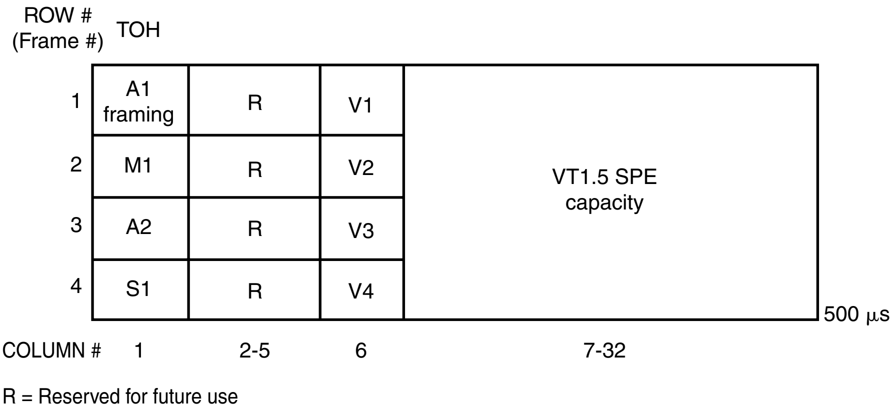


Figure 3-32 VT1.5 interface superframe format.

payload of the previous frame. The count is placed in bits 1 through 4 of the path status byte G1 in the opposite direction of transmission.

- **Loss of pointer-path (LoP-P)** LoP-P is a defect that occurs when either a valid pointer is not detected in eight contiguous frames or when eight contiguous frames are detected with the new data flag (NDF) set to 1001 without a valid concatenation indicator (see ANSI T1.105). An LoP-P is terminated when either a valid pointer with a normal NDF set to 1001 or a valid concatenation indicator is detected for three contiguous frames or the occurrence of an AIS-P defect.
- **AIS-path (AIS-P)** An AIS-P defect is the occurrence of the path AIS in three contiguous frames. An AIS-P defect is terminated when a valid STS pointer is detected with the NDF set to 1001 for a single frame 0110 for three consecutive frames.

  Path AIS is an all-1s signal in H1, H1*, H2, H2*, H3, and the entire SPE.
- **Trace identifier mismatch-path (TIM-P)** When activated, the TIM-P defect needs to be detected within 30 seconds when none of the sampled trace identifier messages match the provisioned expected value. The TIM-P defect will be terminated within 30 seconds when four-fifths of the sampled trace identifier messages match the provisioned expected value.
- **Unequipped-path (UNEQ-P)** A UNEQ-P defect occurs when the C2 bytes in five consecutive frames containing 0000000 are detected. A UNEQ-P defect is terminated when no C2 bytes in five consecutive frames containing 00000000 are detected.
- **RDI-path (RDI-P)** Two types of RDI-P are available: unenhanced (RDI-P) and enhanced (ERDI-P). Only one type can be supported by a PTE at a time, and the enhanced type is preferred. The enhanced type is defined such that it is generally compatible with the RDI-P.
  - **RDI-path (RDI-P)** An RDI-P defect is the occurrence of the RPI-P signal in 10 contiguous frames. An RDI-P defect is terminated when no RDI-P signal is detected in 10 contiguous frames. The setting of RDI-P in response to LoP-P and AIS-P is specified in ANSI T1.231.
  - **ERDI-path (ERDI-P)** An ERDI-P defect is the occurrence of any 1 of 3 ERDI-P signals (codes) in 10 contiguous frames. An ERDI-P defect terminates when no ERDI-P signal is detected in 10 contiguous frames. The setting of ERDI-P in response to LoP-P and AIS-P, TIM-P or UNEQ-P, or PLM-P is specified in ANSI T1.231.

- **Path label mismatch-path (PLM-P)** This occurs when the C2 bytes in five consecutive frames contain a payload label different than the labels allowed by the receiver. A PLM-P defect is terminated when no C2 bytes in five consecutive frames contain a payload label different than a label allowed by the receiver. PLM-P is not a path-level defect, but does cause the setting of one of the ERDI-P codes. In addition, this code is also used in ATM services to indicate a loss of cell delineation defect.

Performance and Failure Alarm Monitoring Performance and failure alarm monitoring requirements for the physical media, section, line, and path levels are specified in ANSI T1.231.

Performance-monitoring functions in terms of collection, storage, validity, register size, threshold and alerting, reporting, accuracy, and resolution are the same as for SONET path—that is, errored second path.

3.4 Reconfiguration Support

This section focuses on the reconfiguration capabilities of SONET-based systems. The section is based on Nortel's materials, which are used with permission.[13] For illustrative purposes, some of the features of the product line are briefly highlighted.

3.4.1 SONET Networking Topologies

Today's SONET transport networks typically employ a number of different topologies to satisfy important objectives for network simplicity, cost containment, bandwidth efficiency, and survivability. For example, an optical hubbing configuration may be used to eliminate the need for a costly and complicated arrangement consisting of several back-to-back network elements. Similarly, a self-healing ring can be deployed to ensure service survivability through redundant, geographically diverse paths. This section defines each of the major SONET network element configurations and discusses their key attributes.

Point-to-Point Terminal The point-to-point terminal configuration is the classic transmission topology that terminates the entire SONET payload at each end of a fiber span or route. Point-to-point systems are typically employed in a basic transport application calling for a single-system/single-route solution. Although not specifically designed to be completely survivable, the reliability of a point-to-point system may be enhanced through a geographically diverse protection path (see Figure 3-33). Where diverse routing exceeds the normal reach of the transport system, one or more regenerators or optical amplifiers can be deployed to reconstitute (or boost) the optical signal.

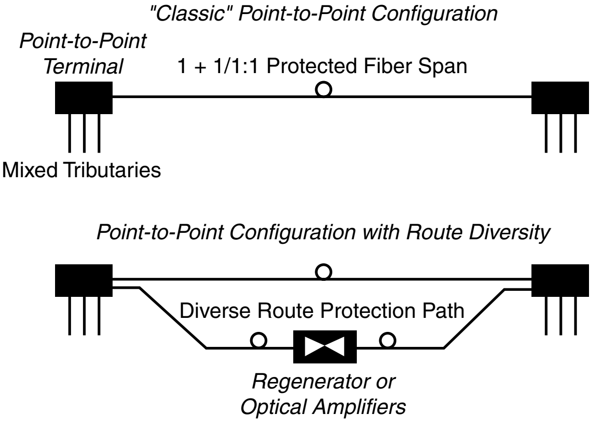
1:N Protection Channel Sharing This is a multinetwork element point-to-point configuration that conserves fiber pairs by enabling multiple systems to share a common protection channel. In applications with rapidly growing traffic demand, the 1:N protection arrangement can defer, or avoid altogether, the large capital outlays and long lead times associated with the deployment of new fiber cable.

If a failure or degradation occurs on any working optical channel, the affected traffic is automatically rerouted over the protection channel via an intershelf protection loop and a dedicated protection shelf at each end of the span (see Figure 3-34).[14] Each working shelf can be equipped with a full complement of mixed electrical and optical tributary types.

If desired, the 1:N protection configuration's protection shelves and the normally empty protection bandwidth may be exploited to transport unprotected extra traffic.

Linear ADM Linear ADM configuration provides direct access to individual eastbound or westbound VT/STS channels at intermediate sites along a fiber route—without the unnecessary multiplexing/demultiplexing

Figure 3-33
Point-to-point systems with and without router diversity.



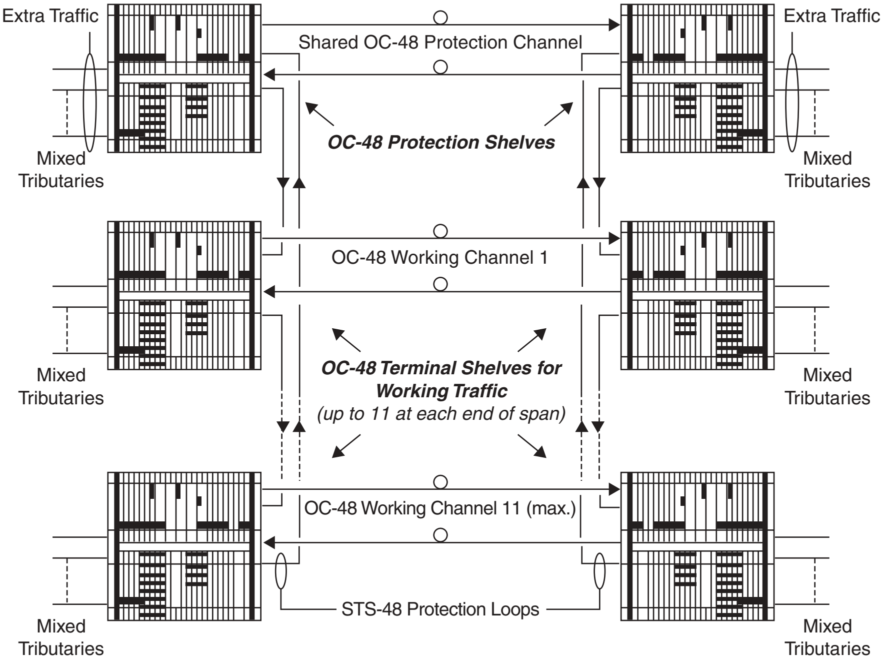


Figure 3-34
1:N protection channel sharing configuration—Nortel's S/DMS TransportNode OC-48 application is shown.

of pass-through traffic (see Figure 3-35). It offers cost savings and improved reliability in comparison to intermediate sites equipped with complex back-to-back terminal arrangements.

As with point-to-point terminal applications, the survivability of linear add/drop routes can be enhanced through geographically diverse protection paths. Alternatively, a high degree of survivability may be achieved by using a hybrid linear ADM subtending ring configuration known as a *path-protected linear ADM route*. Subtending rings are covered in the following section.

As an example, Nortel's S/DMS TransportNode's linear ADMs support mixed electrical/optical tributary types and flexible VT/STS time slot assignment technology. The latter enables any desired routing of any tributary VT/STS channel to any eastbound or westbound VT/STS channel—or direct pass-through from east to west—as desired for optimum use of fiber span capacity and tributary hardware.

An Introduction to Self-Healing SONET Rings Self-healing ring architectures (see Figure 3-36) are the preferred solution for applications where survivability is of the utmost importance. They protect against cable

Figure 3-35 Linear add/drop route example.

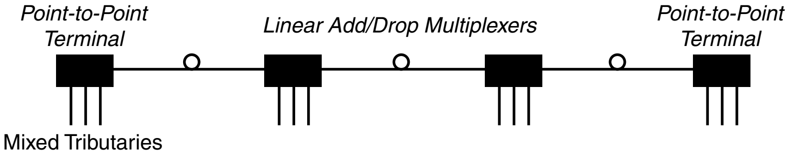
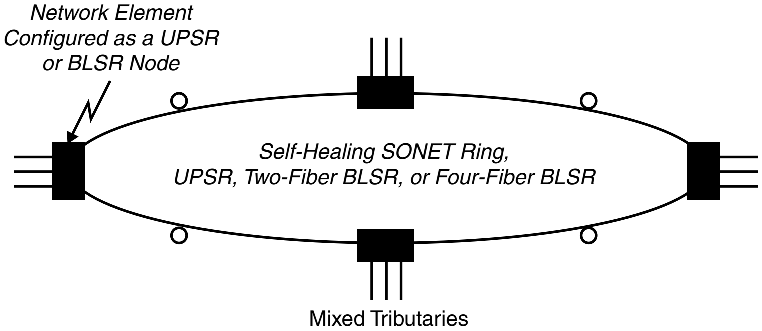


Figure 3-36 Ring configuration employed by all types of SONET self-healing rings.



cuts and node failures by providing duplicate, geographically diverse paths for each service. To ensure end-to-end survivability for services traversing multiple rings, adjacent rings may be interlocked using redundant matched-node inter-ring gateways.

A SONET ring is generally of one of three types: a UPSR, two-fiber BLSR, or four-fiber BLSR. Although each type of ring features fully self-healing operation, differing characteristics may make one architecture preferable in certain situations. For example, a two-fiber BLSR is a good choice for networks with a highly distributed mesh traffic pattern. Table 3-3 shows ideal applications for each SONET ring architecture.

Subtending Rings A subtending ring is an advanced dual-ring configuration where a node's tributary optics support a secondary ring (see Figure 3-37). This arrangement offers substantial capital savings and network simplification at a CO hub by allowing a single network element to serve in place of multiple collocated shelves. A subtending ring can employ either UPSR or BLSR technology.

Subtending rings may be arranged for single homing (as in Figure 3-37) or, for greater geographic diversity, in a dual-homing architecture where the primary and subtending rings intersect at two points (see Figure 3-38). Subtending rings may also be deployed with ADM in a hybrid configuration to provide survivability for linear ADM routes. This is a path-protected linear ADM route (see Figure 3-39).

Table 3-3

Ring Architecture Comparison

Ring Architecture	Best Used In	How It Works
UPSR	Access networks where most traffic terminates in a CO hub	See Section 3.4.2.
Two-fiber BLSR	Access and interoffice networks with a highly distributed traffic pattern	See Section 3.4.3.
Four-fiber BLSR	Applications requiring extra-high capacity and/or protection against multiple concurrent faults	See Section 3.4.4.

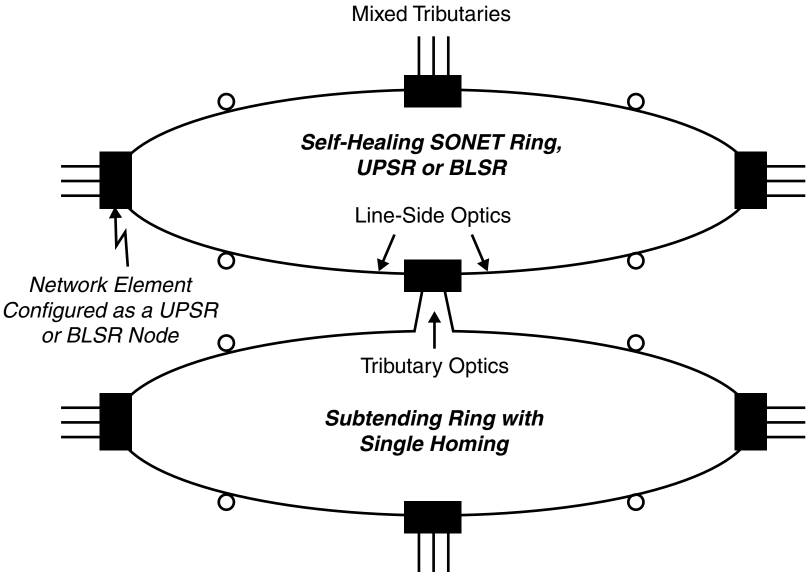


Figure 3-37
Subtending ring with single homing.

As an example, Nortel's S/DMS TransportNode OC-3 Express network elements currently support subtending UPSRs with either single or dual homing. In addition, a subtending BLSR feature is planned for future introduction on S/DMS TransportNode OC-192 systems.

Folded Rings In a configuration known as a folded ring, service providers can implement a UPSR or BLSR even where existing fiber infrastructure does not support a true route diversity ring system. The folded ring functions in exactly the same manner as a UPSR or BLSR, but the fibers on each side of the ring share a common conduit, as shown in Figure 3-40. This solution enables a service provider to begin the conversion process from a

Figure 3-38
Subtending ring with dual homing.

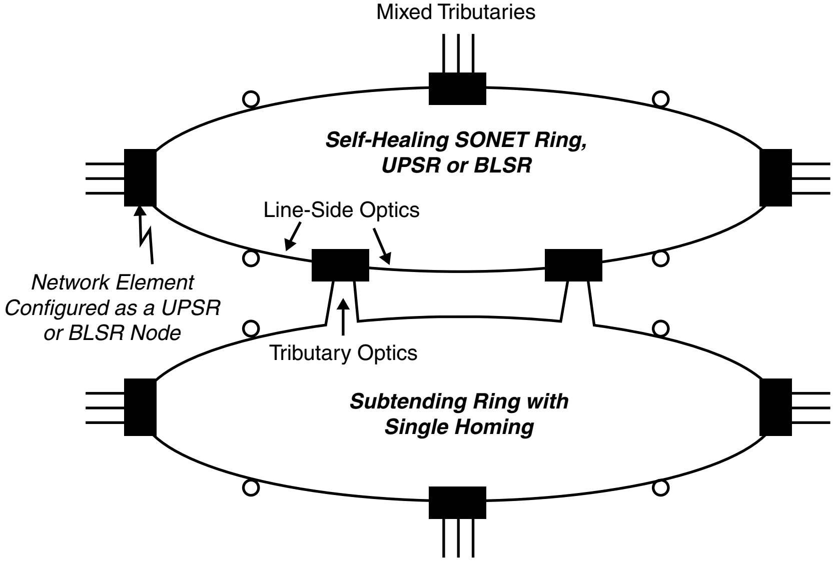


Figure 3-39
Path-protected linear ADM route formed by interlocked subtending and primary rings.

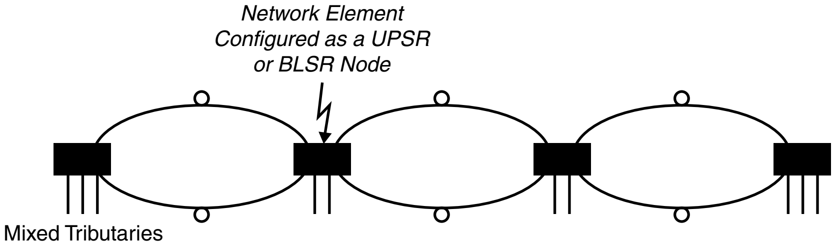
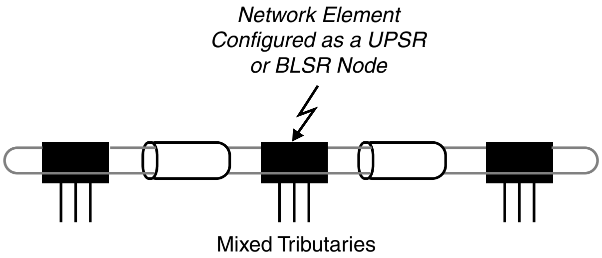


Figure 3-40
Nonroute diverse folded ring.



linear to a ring system today and then easily upgrade to a route diversity ring as additional fiber routes are constructed.

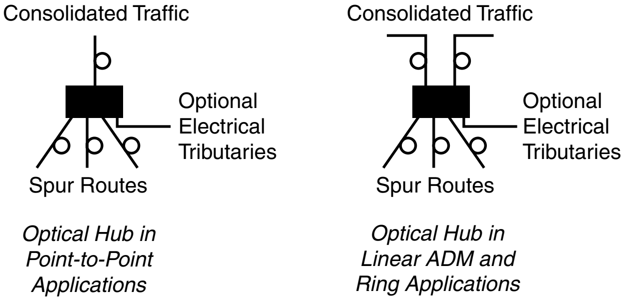
Optical Hubbing Optical hubs consolidate traffic from multiple spur routes onto an optical channel extending to a remote site. The configuration eliminates the cost and complexity of multishelf arrangements that would otherwise be necessary to pass traffic from several lower-rate fibers to a single higher-rate system. The advantages of optical hubbing apply equally to point-to-point, linear add/drop, and ring systems (see Figure 3-41).

As an example, Nortel's S/DMS TransportNode OC3/12/48/192 network elements all support optical tributaries for hubbing applications. For added flexibility, mixed electrical/optical tributary interfaces may be combined on the same shelf as desired. OC-3 Express network elements offer the additional advantage of tributary-to-tributary hairpinning for the convenient local termination of traffic entering from any spur route. (This feature was planned for introduction on OC-192 systems as well.)

Regenerators A regenerator extends system reach by reconstituting the optical signal at an intermediate point between two service-terminating locations. If necessary, multiple cascaded regenerators can be deployed to reach hundreds of miles (see Figure 3-42). Unlike some optical amplifier solutions, regenerators interwork with the SONET overhead section bytes for greater flexibility in operations access and improved isolation of troubles to a specific section along a fiber route.

Regenerators are generally of one of two types. A *nonroute diverse* regenerator supports two bidirectional optical channels in each direction to

Figure 3-41
Network elements configured as optical hubs.



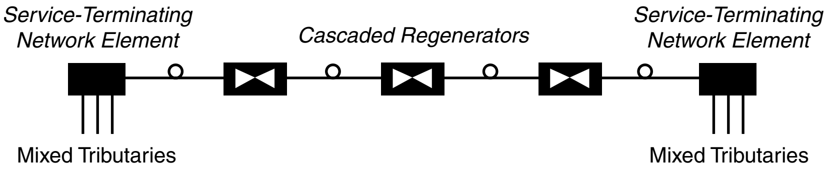


Figure 3-42 Extended reach application supported by cascaded regenerators.

reconstitute the optical signal on both the working and protection fibers of a 1+1/1:1 protected point-to-point or linear system without route diversity. A *diverse route regenerator* supports the single bidirectional optical channel associated with two-fiber BLSRs and other route diversity arrangements (such as the example depicted in Figure 3-33).

3.4.2 Understanding OC-3 UPSRs

A UPSR is a survivable, closed-loop transport architecture that protects against cable cuts and node failures by providing duplicate, geographically diverse paths for each service. Adjacent nodes on the ring are interconnected using a single pair of optical fibers. UPSRs feature a unique unidirectional traffic flow, as illustrated in Figure 3-43. Working traffic travels in only one direction (clockwise) on the ring, whereas a second unidirectional protection path is provided in the counterclockwise direction.

OC-3 UPSRs are typically deployed in access networks where traffic from multiple carrier serving areas (CSAs) terminates at a CO hub site. (BLSRs are often the preferred choice for other types of traffic.) Because the UPSR supports virtual ring configurations, isolated CSAs/nodes can be easily placed on the same ring with several other physically interconnected nodes.

Some advantages are exclusive to the UPSR architecture. One such advantage is that services may originate and terminate on the same UPSR or may be passed to an adjacent access or interoffice ring for survivable transport to the service-terminating location. OC-3 UPSRs support matched-node inter-ring gateways that ensure survivable interconnections between rings irrespective of the type (UPSR or BLSR), line rate, or equipment supplier.

UPSR Operation All traffic is managed at the path (service) layer on an OC-3 UPSR. As each individual DS1, DS3, or STS-1 service travels in one direction around the ring, a duplicate signal passes in the opposite direction for protection (see Figure 3-44). A path selector continuously

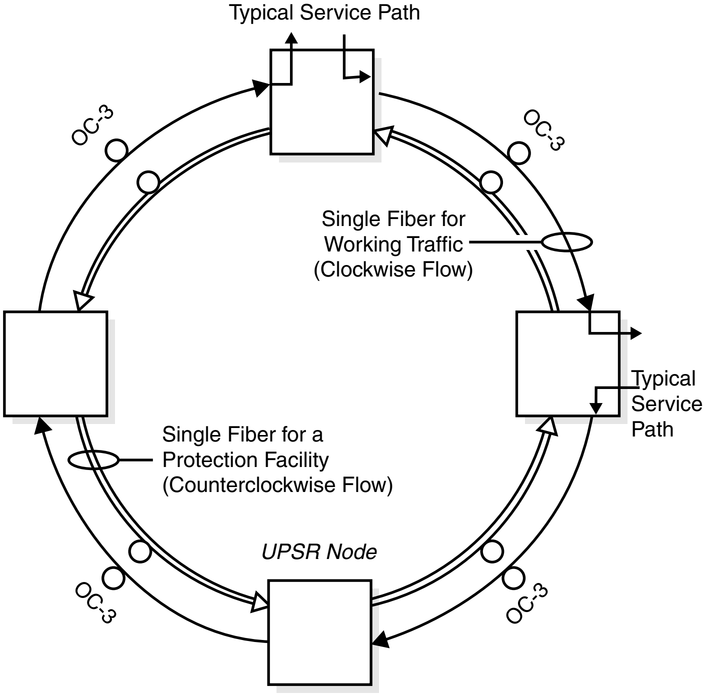


Figure 3-43
Simplified view of OC-3 UPSR architecture.

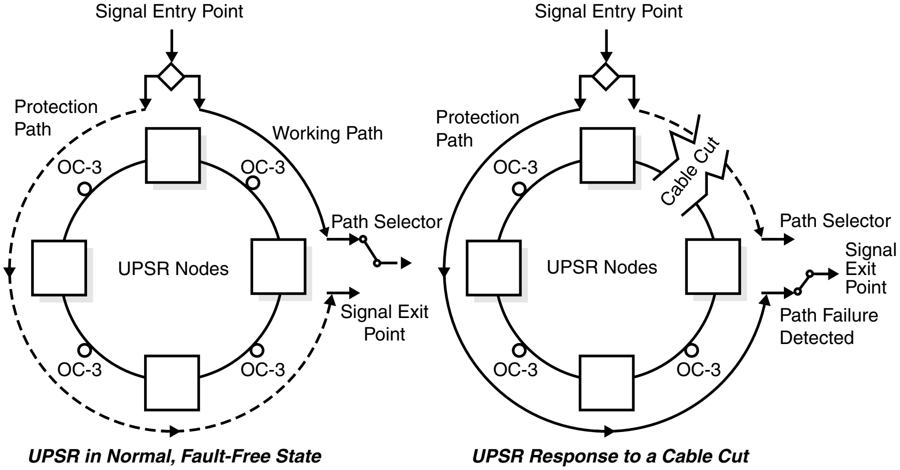


Figure 3-44
OC-3 UPSR protection switching—one direction of transmission is shown.

monitors both the working and protection signals at each end of the path and automatically switches to the protection signal in the event of optical span or node failure or degradation. The path selector detects path failure and signal degradation based on industry-standard parameters such as the following:

- Path AIS
- Path LoP
- Signal degrade (SD)
- Excessive path-layer BIP errors

Once protection switching has been activated, traffic flows in a traditional bidirectional pattern that is similar to a linear add/drop topology. The S/DMS TransportNode OC-3 Express product line supports UPSR protection switching in accordance with the GR-1400 industry standard.

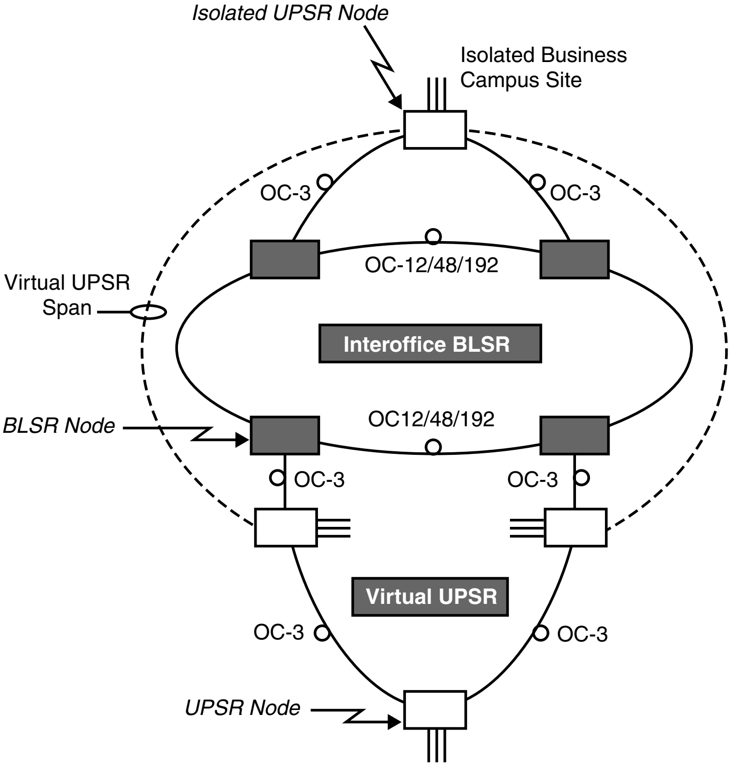
Support for Virtual Rings Because intermediate nodes are essentially transparent in the UPSR protection-switching scheme, other survivability architectures such as a two-fiber or four-fiber BLSR may be incorporated within a UPSR in a hybrid arrangement known as a *virtual ring* or *path-in-line* ring. This alternative is an excellent choice for interconnecting an isolated business campus or CSA with several remote sites that share a common access network, as shown in Figure 3-45. Virtual rings can also provide a survivable, geographically diverse path to an interexchange carrier (IXC) point of presence (POP).

Path protection switching in a virtual ring operates in exactly the same manner as a conventional UPSR since any intermediate BLSR nodes (or other foreign nodes) are completely transparent to the virtual ring. Similarly, an intermediate BLSR protects its portion of the path using standard line-layer BLSR ring protection switching (RPS). Note that survivability is completely ensured without use of protected optics on the OC-3 links between the virtual ring and an intermediate BLSR. This advantage can yield substantial circuit pack inventory savings and frees up shelf space for other revenue-generating services.

3.4.3 Understanding Two-Fiber BLSRs

As with the OC-3 UPSR discussed previously, the two-fiber BLSR is a survivable SONET transport architecture that protects against cable cuts and node failures by providing duplicate, geographically diverse paths for each service. With the capability to reuse bandwidth as traffic is added/dropped at various locations around the ring, two-fiber BLSRs are ideal for distributed mesh and node-to-adjacent-node traffic patterns typical of interoffice networks and are sometimes found in access networks as well.

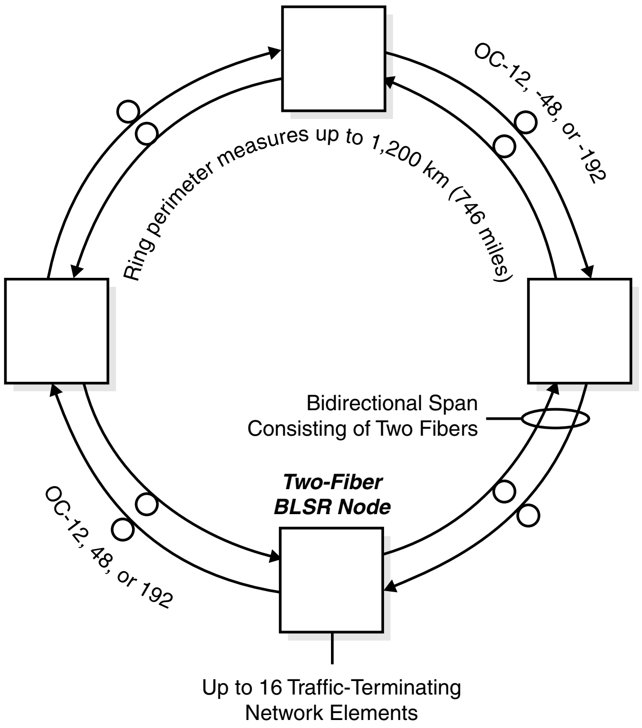
Figure 3-45
Virtual OC-3 UPSR architecture.



Nortel's S/DMS TransportNode offers two-fiber BLSRs operating at a choice of line rates: OC-12, OC-48, or OC-192. All S/DMS TransportNode two-fiber BLSRs comply with GR-1230 industry standards and support matched-node inter-ring gateways to ensure survivability for services traversing multiple interconnected rings. These gateways are fully interoperable with two-fiber BLSRs, four-fiber BLSRs, and UPSRs irrespective of line rate or equipment vendor. Nortel's S/DMS TransportNode also offers BLSR configurations supporting extra traffic—a feature that enhances the revenue potential of existing fiber plant.

The Two-Fiber BLSR Architecture In the two-fiber BLSR architecture, SONET OC-12/48/192 network elements are interconnected in a closed loop using bidirectional spans consisting of two fibers (see Figure 3-46). Each fiber handles one direction of transmission similar to traditional point-to-point and linear add/drop topologies. The perimeter of the ring can measure up to 1,200 km (746 miles) and up to 16 traffic-terminating nodes can be

Figure 3-46 Simplified view of a two-fiber BLSR architecture.



placed on the ring. Regenerators or optical amplifiers may be deployed between nodes as needed to maintain the level of the optical signal.

Allocation of Bandwidth Exactly one-half of the bandwidth available between adjacent nodes can be used for working traffic, and the remaining bandwidth is reserved for protection. As Figure 3-47 suggests, this bandwidth partitioning is accomplished by mapping traffic into STS-1 channels (time slots) 1 through 6 in OC-12 applications, into channels 1 through 24 in OC-48 applications, or into channels 1 through 96 in OC-192 applications.

Note that STS-1 time slot usage within a two-fiber BLSR *does not* impose a bandwidth penalty since any service can be routed from either side of the ring (see Figure 3-48). Thus, a two-fiber BLSR node has a maximum fully protected traffic termination capacity equivalent to STS-12 (for OC-12 applications), STS-48 (for OC-48 applications), or STS-192 (for OC-192 applications) bandwidth.

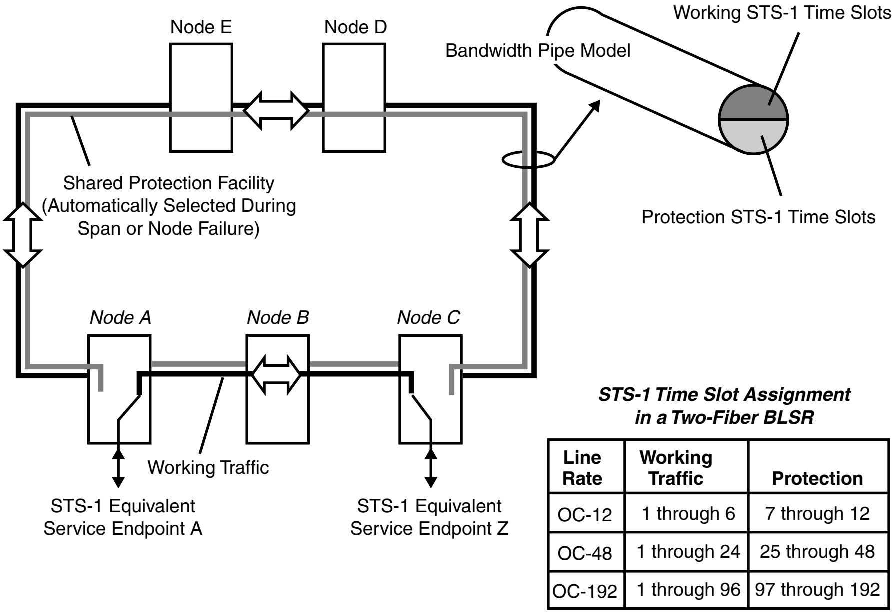


Figure 3-47 Span bandwidth allocation in a two-fiber BLSR.

STS-1 Time Slot Assignment in a Two-Fiber BLSR

Line Rate	Working Traffic	Protection
OC-12	1 through 6	7 through 12
OC-48	1 through 24	25 through 48
OC-192	1 through 96	97 through 192

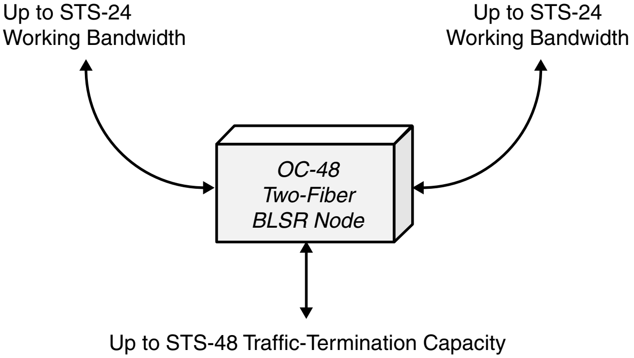


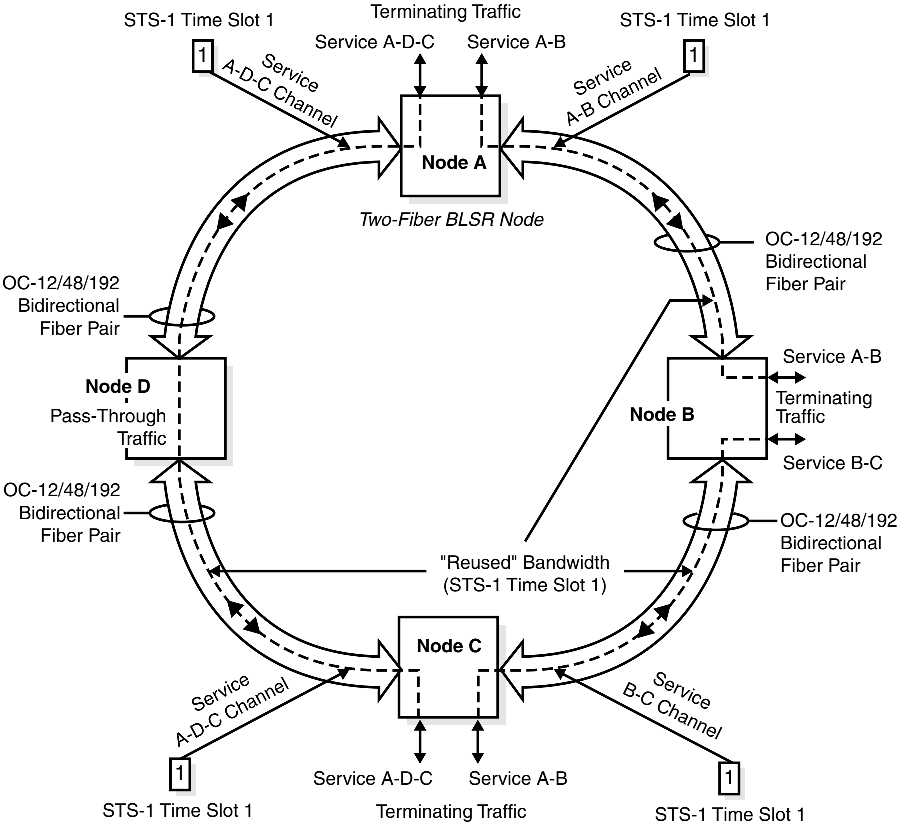
Figure 3-48 Two-fiber BLSR dual-traffic routing—an OC-48 application is shown.

Service Delivery via a Two-Fiber BLSR A service path is provisioned on a two-fiber BLSR by selecting endpoint network elements, tributary ports, and an STS-1 time slot linking the service entry and exit points. A service may reach its destination by traveling in either direction around the ring, as needed for optimum utilization of the available STS-1 channels. Intermediate nodes on the service path, if any, simply pass the service from east to west without modifying the STS-1 channel assignment.

The two-fiber BLSR architecture enables STS-1 channels to be reused as traffic is terminated at various locations around the ring—a feature that makes the architecture ideally suited for the distributed mesh and node-to-adjacent-node traffic patterns of interoffice networks. Reusable bandwidth also offers important synergies in ATM networks.

Figure 3-49 illustrates how bandwidth can be reused on a two-fiber BLSR through application of STS-1 time slot assignment (TSA) technology. Note that Service A-B is routed from Node A to Node B around the east side of the ring using STS-1 channel 1. Because Service A-B terminates at Node B, STS-1 time slot 1 can be reused to transport Service B-C. The same channel is reused again at Node C to transport Service A-D-C. Node D passes Service A-D-C through to Node A without modifying its time slot assignment. Thus, the sample two-fiber BLSR depicted in the figure is able to

Figure 3-49
Reusing bandwidth in a two-fiber BLSR.



transport traffic having STS-3 total bandwidth while using only a single STS-1 channel.

Automatic Healing of Failed or Degraded Optical Spans In the event of failure or degradation in an optical span, automatic RPS reroutes affected traffic away from the fault within 50 ms, preventing a service outage. Traffic is redirected by looping back STS-1 time slots (see Figure 3-50). Logically, the normally unused protection bandwidth bridges the defective span thereby maintaining service for all terminating and pass-through traffic. When setting up the bidirectional loopback path, OC-12 systems map working STS-1 time slots 1 through 6 to protection time slots 7 through 12 (respectively), OC-48 network elements map working STS-1 time slots 1 through 24 to protection time slots 25 through 48, and OC-192

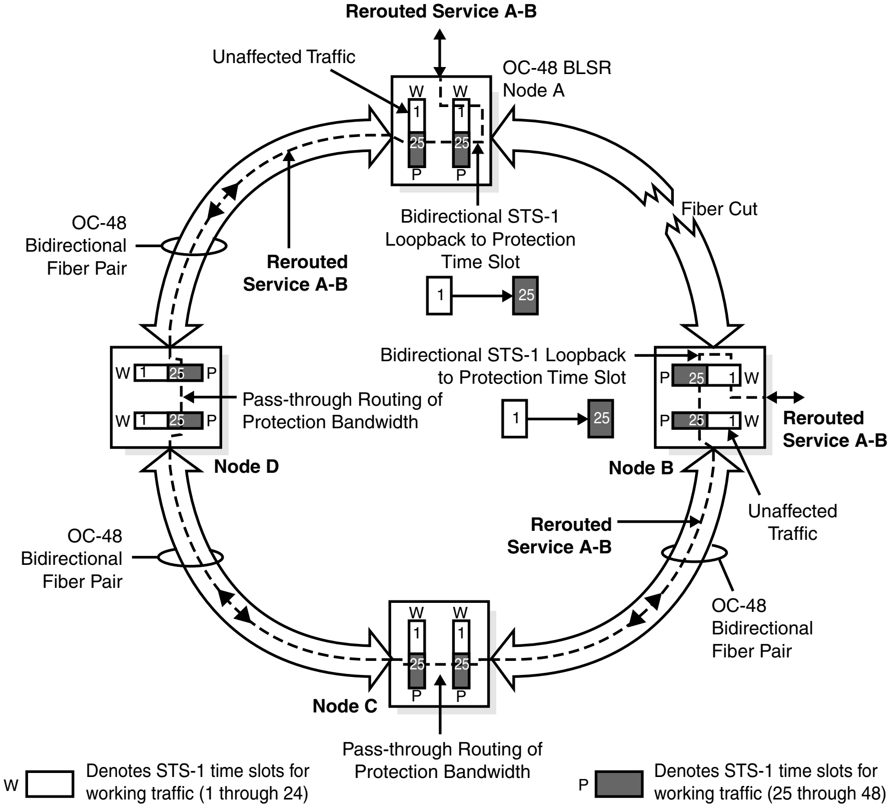


Figure 3-50
Two-fiber BLSR response to an optical span failure—an OC-48 application is shown.

systems map working STS-1 time slots 1 through 96 to protection time slots 97 through 192.

STS-1 time slot bridging occurs only at the nodes adjacent to the fault, with intermediate nodes (for example, Nodes C and D in Figure 3-50) simply passing through the redirected traffic mapped to the protection time slots. STS-1 time slot assignments for working traffic at intermediate nodes are unaffected by the fault.

Conditions that trigger RPS include total failure modes such as LoS and also degradation in terms of excessive line-layer BIP errors. Because protection switching is revertive in a two-fiber BLSR, traffic automatically returns to its normal routing—without human intervention—after a fault-free state exists for a user-defined wait-to-restore interval.

Rerouting of Pass-Through Traffic During Node Failures The two-fiber BLSR architecture also fully protects all restorable traffic in the event of a node failure anywhere along the ring. Although tributaries terminating at the failed node cannot be protected, traffic passing through that node is automatically redirected away from the fault via a time slot loopback similar to the previous span failure example of Figure 3-50. In an action referred to as *squelching*, nodes adjacent to the failure replace nonrestorable traffic with a path-layer AIS to notify the far end of the interruption in service. The squelching feature employs automatically generated squelch maps that do not require manual record keeping.

Capacity Advantages of Two-Fiber BLSRs Due to the BLSR's capability to reuse STS-1 channels, a BLSR may offer significant capacity advantages over UPSRs, depending on the traffic pattern (see Figure 3-51). Where traffic is entirely hubbed (as in most access networks), capacity *equals* that of a UPSR operating at the same line rate. As traffic becomes more distributed (or meshlike), however, a BLSR provides a substantial capacity benefit that can typically range up to 300 percent of the capacity of a UPSR with an identical optical rate. The greatest capacity advantage occurs with the node-to-adjacent-node pattern because this traffic model permits a very high level of channel reuse. See Figure 3-52 for a capacity comparison of all three traffic patterns relative to a UPSR.

In a given ring application, the capacity required on a BLSR depends only on the aggregate eastbound or westbound traffic at the busiest node. When implementing the same ring as a UPSR, the aggregate of all services transported by the entire ring must be taken into account.

Synergies with ATM ATM is a transport technology that evolved in the 1990s that can be used for multimedia and data services. ATM offers band-

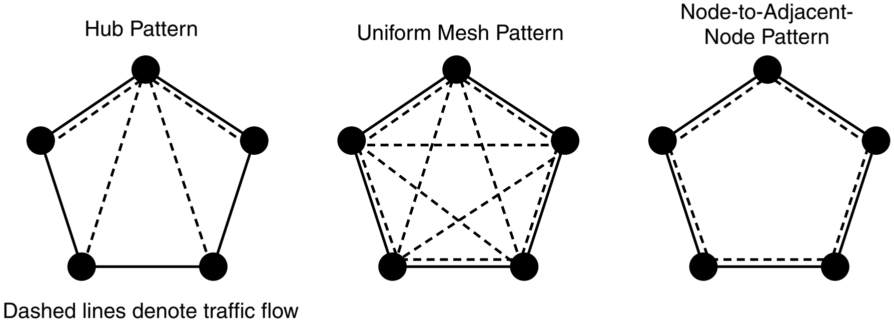


Figure 3-51
Ring traffic patterns.

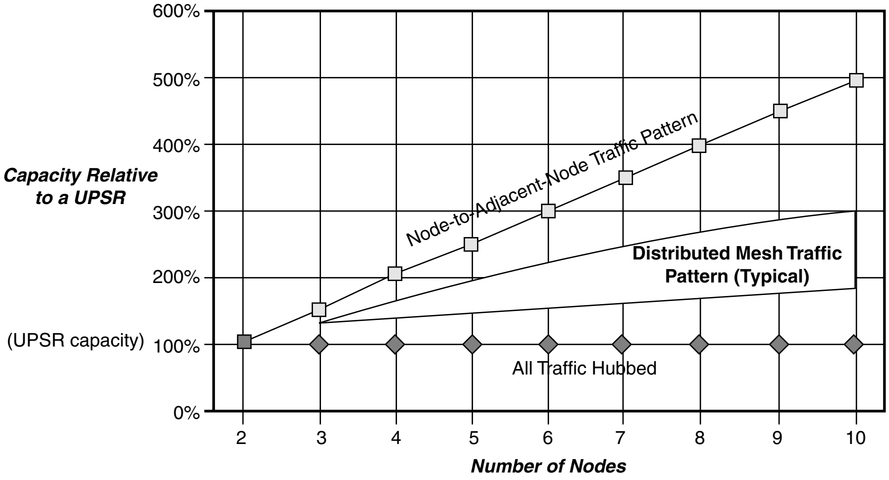


Figure 3-52
BLSR capacity analysis.

width efficiency, scalability, and service transparency for data, voice, video, and multimedia applications. To date, ATM has been extensively deployed in enterprise data networks, especially WANs with large amounts of bursty traffic. In these environments, the statistical multiplexing capabilities of ATM network elements allow the efficient transport of bursty data without unacceptable underutilization of facilities during off-peak periods.

When distributing data embedded in ATM payloads among various WAN sites, two-fiber BLSRs offer two major advantages:

- A level of survivability that meets the most demanding reliability objects for mission-critical applications
- Efficient sharing of a common STS-3c ATM payload among multiple locations

The latter benefit is illustrated in the example ATM network that appears in Figure 3-53. As shown, a two-fiber BLSR links each site in the network. The ring's capability to reuse bandwidth is employed to deliver the same STS-3c concatenated SONET payload (STS-3c Payload A) to each location. This traffic routing enables the ATM network element at each site to efficiently pack traffic into a common payload through statistical multiplexing. Thus, one economical STS-3c channel can replace an expensive arrangement of dedicated channels linking each site with every other site.

Transporting Extra Traffic on BLSRs If desired, the protection bandwidth on a two-fiber (or four-fiber) BLSR may be used to transport unprotected extra traffic. The additional capacity permits the introduction of new revenue-enhancing services without further investment in fiber plant.

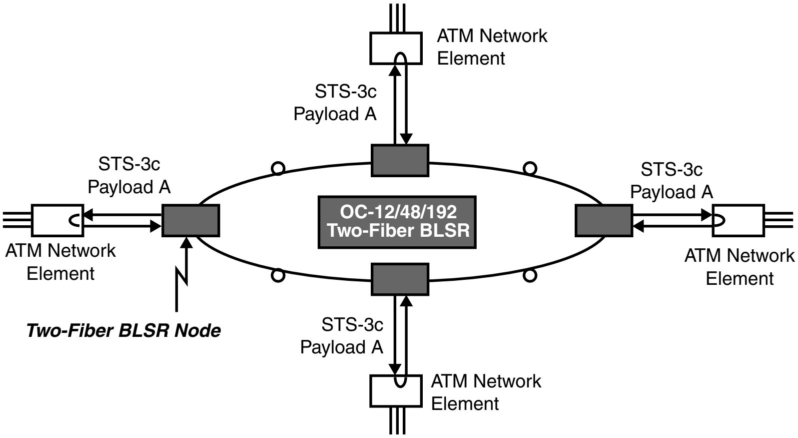
As an example, Nortel's S/DMS TransportNode OC-48 BLSRs offer extra traffic features with an equivalent OC-192 offering now coming to market. Extra traffic can be terminated along with regular protected traffic using any available tributary ports on the network element. If all tributary capacity is exhausted, an additional collocated BLSR node may be deployed to provide additional tributary interfaces as required.

Extra traffic is automatically removed from the protection channels when protection switching occurs on the BLSR.

3.4.4 Understanding Four-Fiber BLSRs

The four-fiber BLSR architecture ensures service survivability through duplicate, geographically diverse paths similar to a two-fiber BLSR. It dif-

Figure 3-53
ATM network interconnected by a two-fiber BLSR.



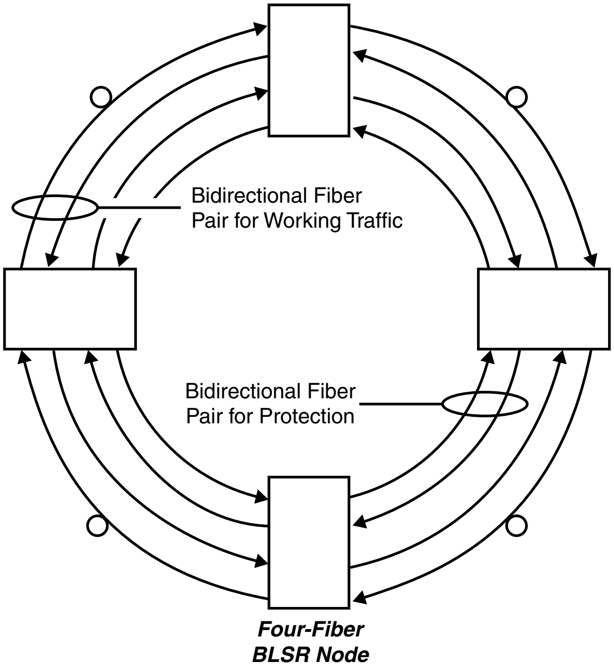
fers from the two-fiber configuration in that four fibers (or two bidirectional fiber pairs) link adjacent nodes on the ring, as shown in Figure 3-54. The additional fiber enhances the architecture in two important ways:

- A doubling in the traffic-handling capacity of the ring since twice the number of fibers are available
- Two protection-switching modes—automatic RPS and traditional 1+1/1:1 span switching for self-healing operations during multiple fault conditions

The allocation of bandwidth in a four-fiber BLSR is identical to point-to-point and linear add/drop topologies. One fiber pair between nodes carries working traffic exclusively, while the other pair serves as a protection facility.

Capacity Advantages of Four-Fiber BLSRs As in a two-fiber BLSR, a node on a four-fiber BLSR may terminate traffic fed from either side of the ring. With a fiber pair in each direction entirely dedicated to working traffic, an individual node on a four-fiber BLSR has twice the traffic-terminating capacity of a two-fiber BLSR node operating at the same line rate. For example, at the OC-192 rate, a four-fiber BLSR node terminates

Figure 3-54 Simplified view of four-fiber BLSR architecture.



the equivalent of up to STS-384 fully protected tributary traffic. In a two-fiber configuration, fully protected tributary traffic is limited to a maximum equivalent bandwidth of STS-192. (This comparison assumes a sufficient number of tributary ports as in Nortel's S/DMS TransportNode OC-192 systems.)

Similarly, the maximum levels of pass-through traffic and unprotected extra traffic terminations are doubled in a four-fiber BLSR. Refer to Table 3-4 for a concise listing of the maximum capacities offered by OC-192 two-fiber and four-fiber BLSRs.

Span Protection Switching Four-fiber BLSR 1+1/1:1 span protection switching bypasses unidirectional or bidirectional faults that affect only the working fiber pair. Examples include a defective splice, faulty connector, and optical transmitter/receiver problems. A cable cut also affects only the working fiber pair in applications where geographically diverse routes link adjacent nodes. Because span protection switching operates independently on each link, service is maintained in the presence of several concurrent working path faults. Thus, in a representative dual-fault scenario (see Figure 3-55), span protection switching can be activated between Nodes A and B to address an optics module failure, while at the same time the protection fibers between Nodes C and D are in use to bypass a defective splice.

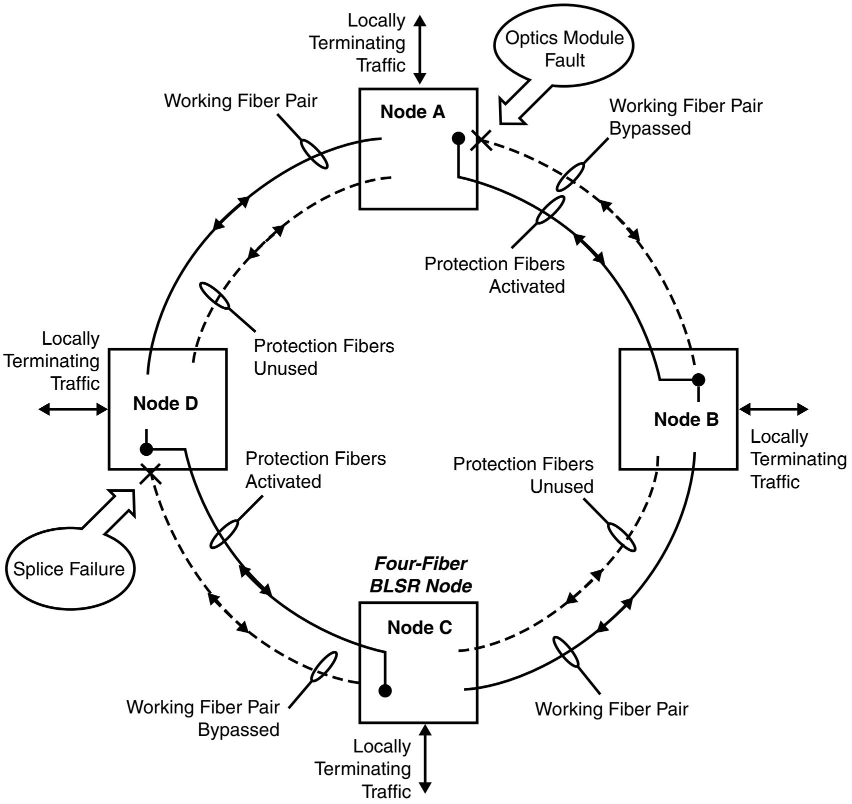
Four-Fiber BLSR Ring Protection Switching (RPS) If a fault affects both the working and protection fibers (a node failure or a cable cut in a span without route diversity), automatic RPS redirects traffic in a manner similar to a two-fiber BLSR. However, instead of looping back time slots within the same fiber pair as in a two-fiber BLSR, the four-fiber architecture loops back traffic from the working fiber pair to the protection fiber pair (see Figure 3-56). Other aspects of four-fiber BLSR RPS are similar to the operation of a two-fiber BLSR.

Table 3-4

OC-192 Two-Fiber/Four-Fiber BLSR Node Capacity Comparison

BLSR Type	Max. Tributary Capacity for Fully Protected Traffic	Max. Pass-Through Traffic on Working Time Slots	Max. Extra Traffic Terminations
Two-Fiber	STS-192	STS-96	STS-192
Four-Fiber	STS-384	STS-192	STS-384

Figure 3-55
Four-fiber BLSR span switching example.

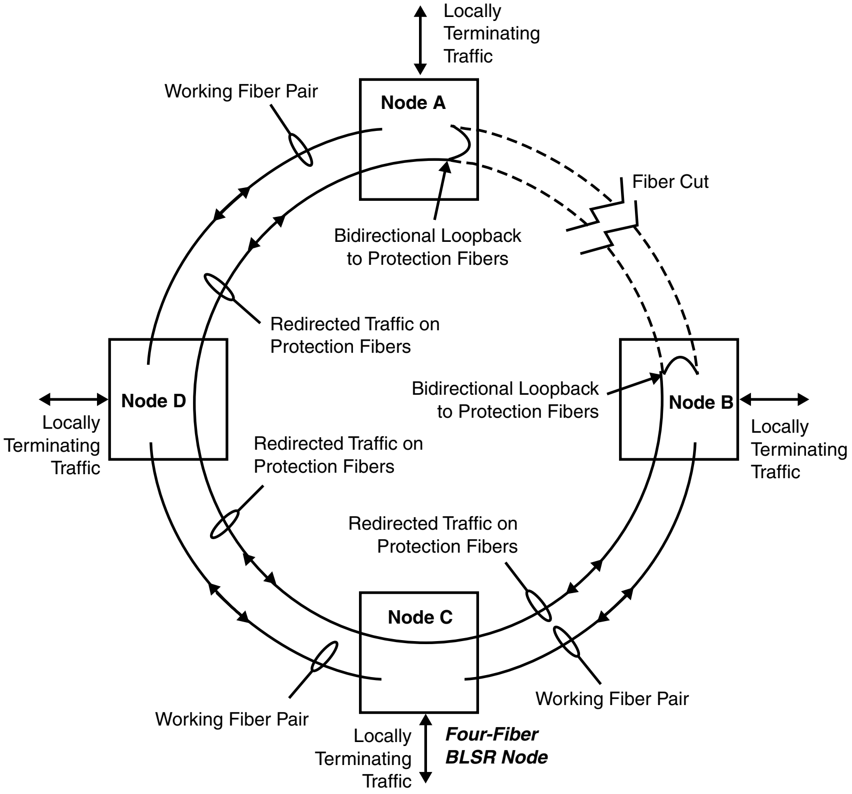


3.4.5 Understanding Matched Nodes

Whereas the route diversity provided by ring topologies fully protects all working traffic passing from node to node along a ring, service paths must be protected on an end-to-end basis. This means survivable ring interconnections are required as traffic passes from an access ring through one or more interoffice rings and then to another access ring serving the service termination point. The SONET matched-node configuration (see Figure 3-57) has been expressly designed to provide these survivable connections between adjacent rings.

Matched-node gateways are completely vendor/technology independent, enabling seamless interworking between rings of different types without regard to line rate or equipment supplier. Thus, matched-node inter-ring

Figure 3-56 Four-fiber BLSR RPS example.



gateways can provide survivable interconnections between an OC-3 UPSR from one equipment supplier and an OC-192 four-fiber BLSR from a different supplier, for example.

Survivable Service Interconnections Between Rings To ensure survivability for services traversing multiple interconnected rings, the matched-node configuration provides redundant routing across inter-ring boundaries as in the BLSR example of Figure 5-58. In the outbound direction, drop-and-continue routing is employed within a primary gateway network element to simultaneously pass an exiting signal to both the neighboring ring and a remote node configured as a secondary inter-ring gateway. For the opposite direction of transmission, the primary inter-ring gateway incorporates a service selector that chooses either the primary or secondary input from the adjacent ring.

Figure 3-57
Inter-ring service transport using survivable matched-node gateways.

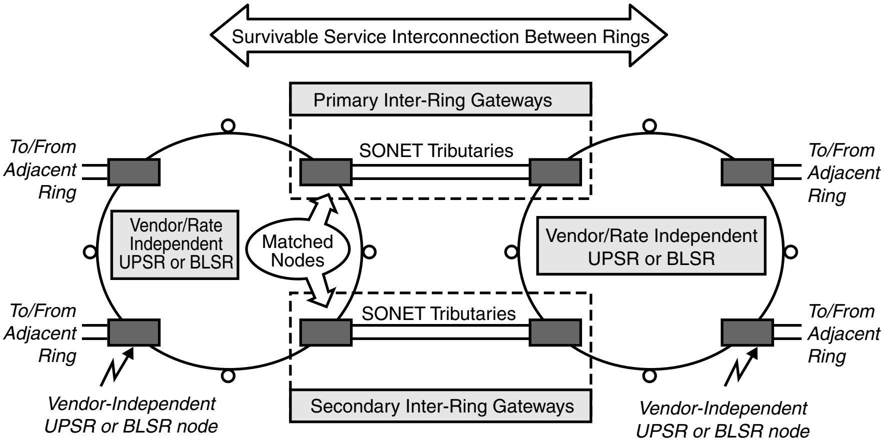
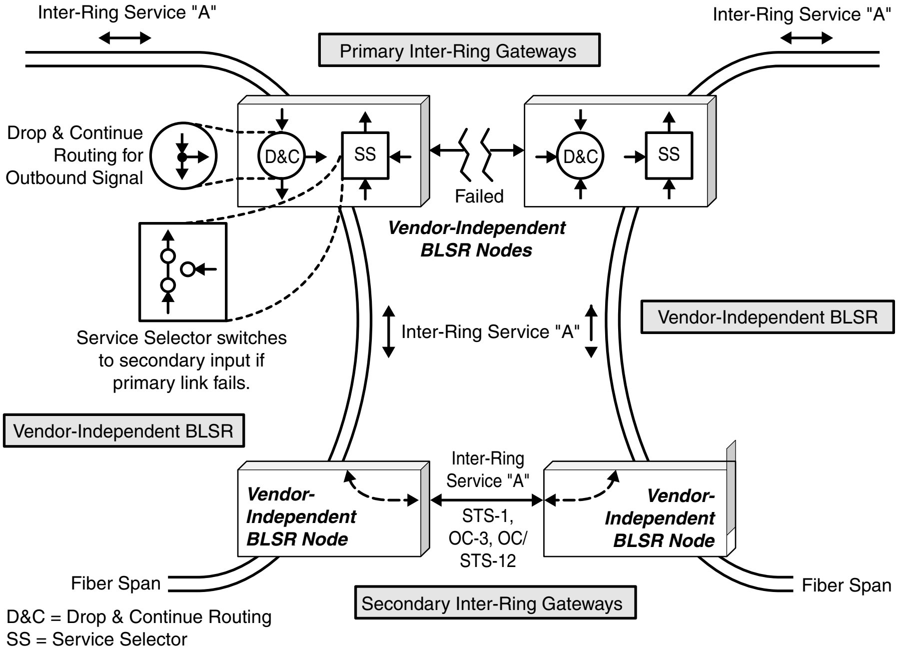
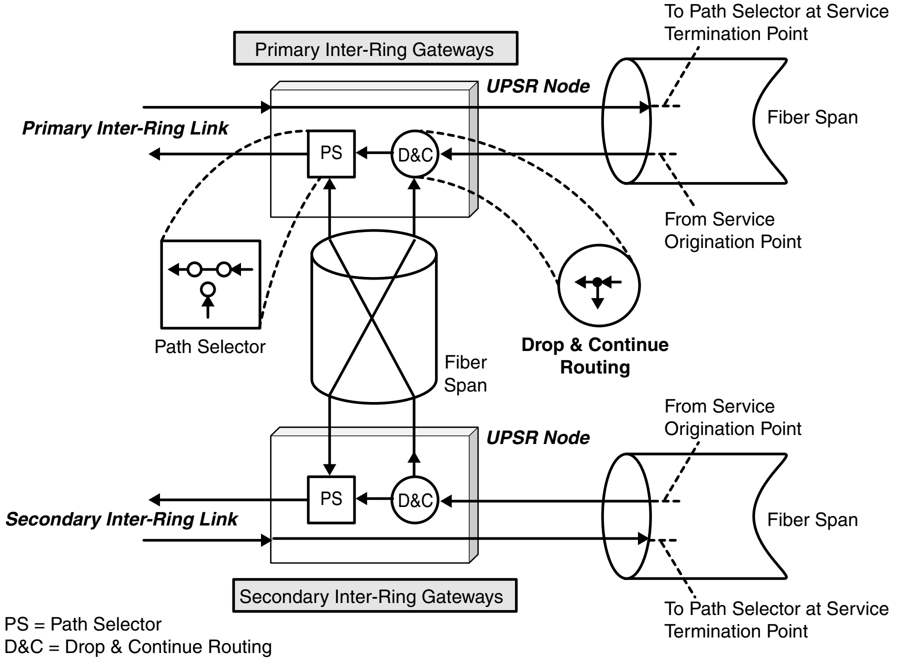


Figure 3-58
Duplicate inter-ring service paths provided by matched-node gateways—a BLSR example is shown.



UPSR matched-node configurations employ standard UPSR path selection techniques plus drop-and-continue routing, as shown in Figure 3-59. From the standpoint of an adjacent ring, this arrangement functions in exactly the same manner as BLSR matched-node inter-ring gateways.

Figure 3-59 UPSR matched-node configuration.



A network element provisioned as a primary gateway for one inter-ring service may be provisioned as a secondary gateway for another inter-ring service as needed for the most efficient use of the ring's available bandwidth. Also, a primary gateway on one ring may feed either a primary or secondary gateway on an adjacent ring as desired.

Adjacent rings are physically interconnected via SONET links (such as STS-1, OC-3, or OC/STS-12 facilities) that serve as conduits for various inter-ring services (such as VT1.5/DS1 or STS-1/DS3 traffic). Routing and selection of inter-ring traffic are always performed at the service layer on both BLSRs and UPSRs.

The service selector or path selector in a BLSR/UPSR matched-node configuration switches from the primary input signal to the secondary input based on standard performance parameters such as line and/or path AIS. In the case of Nortel's S/DMS TransportNode BLSR systems, both line- and path-layer analysis are provided to protect inter-ring services against an extended range of fault conditions compared to other BLSRs that support line analysis only. Path analysis also enables the use of SONET systems on the inter-ring link, permitting, for example, wide separation between interconnecting gateways, as illustrated in Figure 3-60.

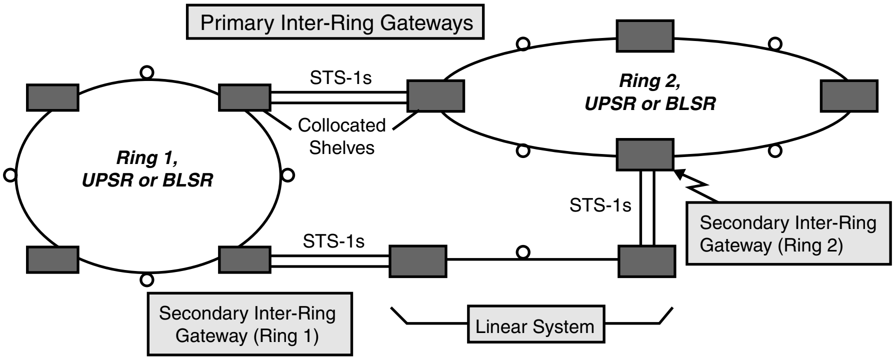


Figure 3-60 Interconnecting remote secondary gateways using a linear SONET system.

For many types of failure situations, end-to-end service is maintained through a combination of RPS and duplicate inter-ring gateways. In the event of a primary gateway failure, for instance, RPS bypasses the faulty node, and inter-ring services pass through the secondary gateway to the adjacent ring.

3.4.6 Wavelength Division Multiplexing (WDM)

WDM multiplies the capacity of existing fiber spans by combining two or more optical signals of different wavelengths on a single fiber. Because it defers—or avoids altogether—the large capital outlays and long lead times associated with deploying new fiber cable, WDM is an ideal solution for critical high-growth routes having an immediate need for more bandwidth. An external coupling device performs the actual mixing of the different optical signals. In *unidirectional WDM*, multiple wavelengths travel in the same direction on an optical fiber, while they pass in opposite directions in *bidirectional WDM* arrangements. Bidirectional WDM is often the preferred approach, especially in applications employing two wavelengths due to the one-to-one association of an individual transport system to an individual optical fiber.

WDM can be further classified as either wideband (also called cross-band), narrowband, or dense depending on the wavelengths involved.

Wideband WDM This WDM technique doubles the capacity of a fiber span by combining the 1,310-nm wavelength with a second wavelength in the low-loss window of optical fiber between 1,528 and 1,560 nm. Both the S/DMS TransportNode OC-12 TBM and OC-48 product lines permit wideband WDM implementations using currently available wavelength options. An example bidirectional wideband WDM configuration based on the TBM's 1,310- and 1,550-nm wavelengths appears in Figure 3-61.

Narrowband WDM As with wideband WDM, the narrowband method provides a twofold increase in fiber span capacity. It employs two low-loss wavelengths, typically 1,533 and 1,557 nm. S/DMS TransportNode's OC-48 and OC-192 product lines both offer wavelength options that fully support bidirectional narrowband WDM, as depicted in Figure 3-62.

On long-haul routes where system reach is an important consideration, narrowband WDM is usually a better choice relative to wideband WDM. In addition, narrowband WDM offers total compatibility with Nortel's S/DMS TransportNode's optical amplifier product line on both single-hop and multihop OC48/192 routes. This means the economic advantages of expanded fiber capacity and extended reach can be achieved in the same application.

Figure 3-61 OC-12 TBM shelves deployed in a bidirectional wideband WDM application.

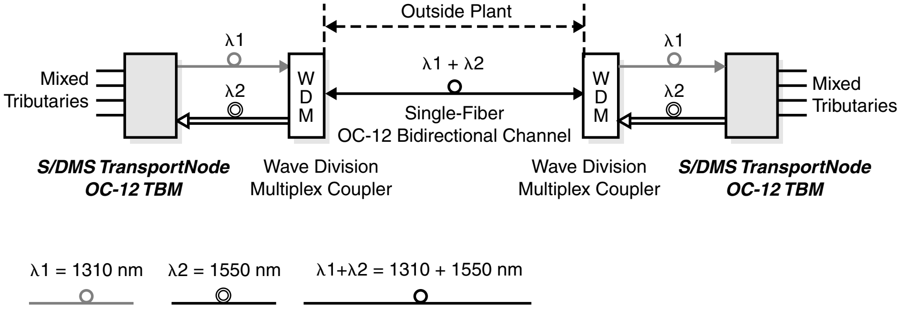
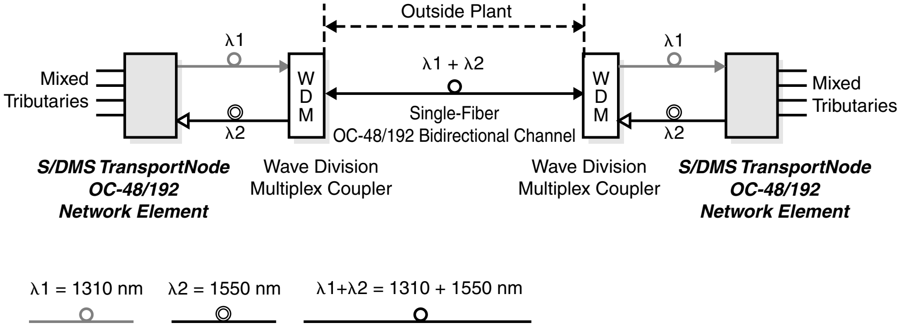


Figure 3-62 Nortel's S/DMS TransportNode OC-48/192 network elements deployed in a typical bidirectional narrowband WDM application.

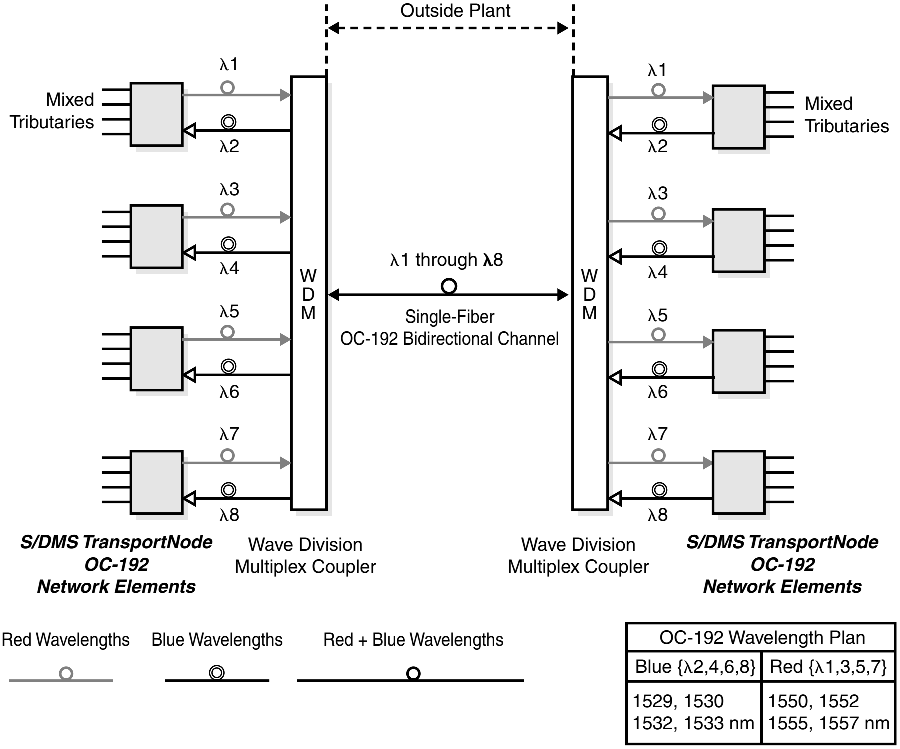


Dense WDM (DWDM) Advanced DWDM technology employs many wavelengths to increase fiber span capacity. These wavelengths fall within two bands: a blue band between 1,529 and 1,541 nm and a red band between 1,549 and 1,557 nm. Each band is dedicated to a particular direction of transmission. Nortel's S/DMS TransportNode offers bidirectional DWDM solutions for OC-48, OC-192, and hybrid OC-48/192 applications. Figure 3-63 illustrates a short-reach bidirectional eight-wavelength OC-192 configuration.

To permit extended reach in DWDM applications, Nortel's S/DMS TransportNode's optical amplifier portfolio includes a leading-edge bidirectional multiwavelength optical repeater (MOR) that handles the associated wavelengths in both the blue and red bands.

This topic is discussed in more detail in Chapter 5.

Figure 3-63
Nortel's S/DMS TransportNode OC-192 bidirectional DWDM configuration (short reach).



OC-192 Wavelength Plan	
Blue {λ2,4,6,8}	Red {λ1,3,5,7}
1529, 1530 1532, 1533 nm	1550, 1552 1555, 1557 nm

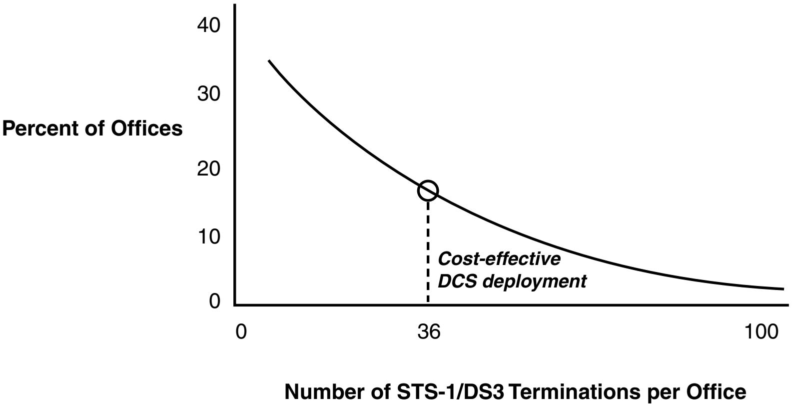
3.4.7 Using SONET Rings to Manage Bandwidth in Small-City Networks

Efficient management of DS1 traffic is a critically important function in just about every interoffice network since most services are transported in the VT/DS1 layer. Effective solutions for traffic routing, grooming, consolidation, and the like must be put in place to avoid escalating operational costs, congestion, and a premature need for network expansion.

Bandwidth management strategies range from manual rearrangement of crossconnect panels in smaller offices to sophisticated DCSs in offices with heavy traffic demand. Although DCSs offer a high degree of flexibility and substantial labor savings, their cost-effectiveness begins to diminish as traffic demand falls below 36 STS-1s/DS3s that require VT/DS1-layer grooming. As can be seen in Figure 3-64, traffic in the vast majority of COs is well below this level.

To manage bandwidth effectively at a reasonable cost in small- to medium-capacity metropolitan networks, many service providers have adopted a crossconnect hub architecture similar to the one shown in Figure 3-65. This approach enables a single DCS at a hub site to handle all bandwidth management needs for an entire network that contains numerous small offices. Traffic requiring grooming and consolidation is first back-hauled to the DCS site before being routed to its final destination. Drawbacks include increased operational complexity and the requirement for additional bandwidth to support back-haul channels.

Figure 3-64
Traffic distribution among COs in typical small- to medium-capacity interoffice networks. Source: Recent Nortel study.



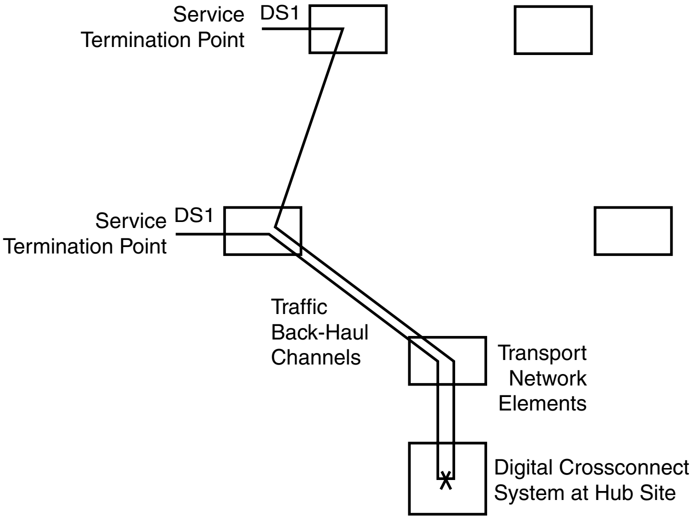


Figure 3-65 Back-hauling traffic to a crossconnect hub.

Bandwidth Management Solutions Based on OC-12 BLSRs Solutions based on VT-managed OC-12 BLSRs offer a bandwidth management alternative for small- to medium-capacity metropolitan networks. A VT-managed TBM ring provides a distributed crossconnect capability that can either enhance an existing crossconnect hub architecture or eliminate the need for an external DCS entirely (see Figure 3-66). The TBM's VT1.5 time slot assignment features groom, consolidate, and route traffic without cumbersome back hauling to a shared DCS. This conserves fiber span capacity as well as DCS ports and core resources. In addition, end-to-end survivability is improved because services don't have to exit the ring for grooming.

Fiber capacity and DCS resource savings stem from two important advantages of the VT-managed BLSR:

- Efficient packing of DS1 traffic into STS-1 channels
- Fewer services back-hauled to a hub site for grooming

These savings can be substantial, depending on the desired network management strategy of the individual service provider. In the small-city traffic demand of Figure 3-67, a VT-managed TBM ring reduces the number of DCS ports required from 14 to 10 for a savings of nearly 30 percent. Note also that this application cannot be handled with a single STS-managed

Figure 3-66
Traffic routing in a VT-managed BLSR.

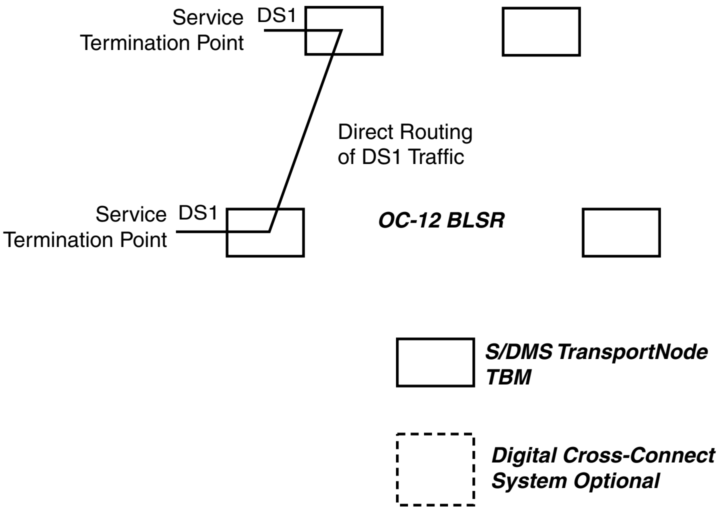
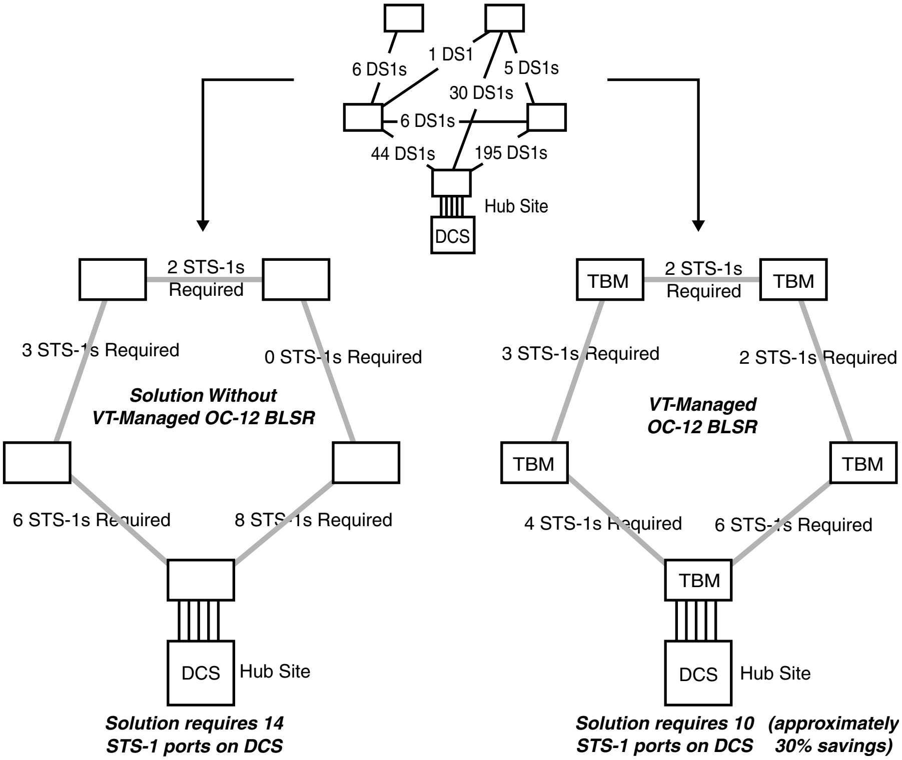


Figure 3-67
Crossconnect savings in a typical small-city network.



OC-12 BLSR since a total of STS-14 working bandwidth would be required at the hub site.

Besides crossconnect port and core resource savings, a VT-managed OC-12 BLSR reduces transport shelf tributary hardware requirements due to the efficient concentration of traffic to a minimum number of DS1 mapper plug-ins. Fiber bandwidth savings increase the effective capacity of the network and thus defer or avoid the need for additional fiber pairs and more network elements. The net result is a substantial reduction in capital costs in the typical small-city network. Also, operational costs and labor are reduced because fewer back-hauled services mean less management complexity.

3.5 Packet over SONET (PoS)

Figure 3-68 depicts the migrations of the broadband data communications protocol stack that have occurred over the past few years or are expected to occur in the near future. In particular, there has been a migration from IP over ATM, IP directly over SONET, or even directly over optics. The goal is to improve efficiency and reduce the cell tax that was associated with the use of ATM.[15] This technology saw deployment in the late 1990s and is currently fairly popular with ISPs. PoS can be used to connect users' routers to the WAN or LAN-to-LAN connections using the SONET/SDH carrier structure.

PoS was originally defined in the IETF RFC 1619, "PPP over SONET and SDH Circuits" (see Table 3-5 for related RFCs) as well as in RFC 1661 and RFC 1662. PoS is a point-to-point solution that eliminates ATM's need for multiple switched virtual circuits. PoS is a Layer 2 transport technology that encapsulates IP packet data directly into SONET. Figure 3-69 is the PoS stack. It clearly enjoys lower overhead than ATM. Users see a 20 to 30 percent increase in capacity when ATM is removed. PoS preserves the existing investments in SONET infrastructure while enabling the deployment of IP-based video and voice applications and other broadband applications. Figure 3-70 depicts a typical application of PoS technology. Figure 3-71 depicts the PoS encapsulation process.

PoS provides limited quality of service (QoS) to IP without the overhead of ATM and has faster recovery from failure than Ethernet-based systems (compared to both the Spanning Tree Algorithm and Rapid Spanning Tree Algorithm). Some see PoS as an interim technology for moving from ATM. The next step is transporting packets directly on the physical layer (IP over

Figure 3-68 Data communications protocol stacks.

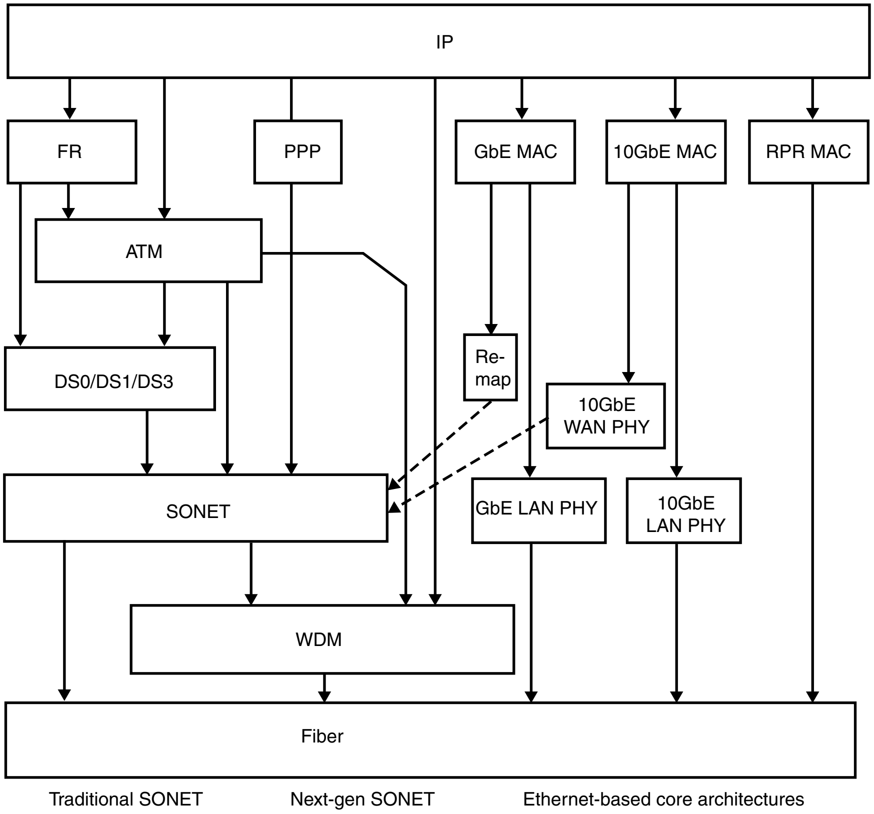


Figure 3-69 PoS stack.

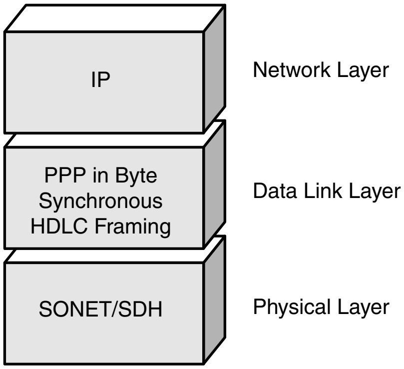


Figure 3-70
PoS example.

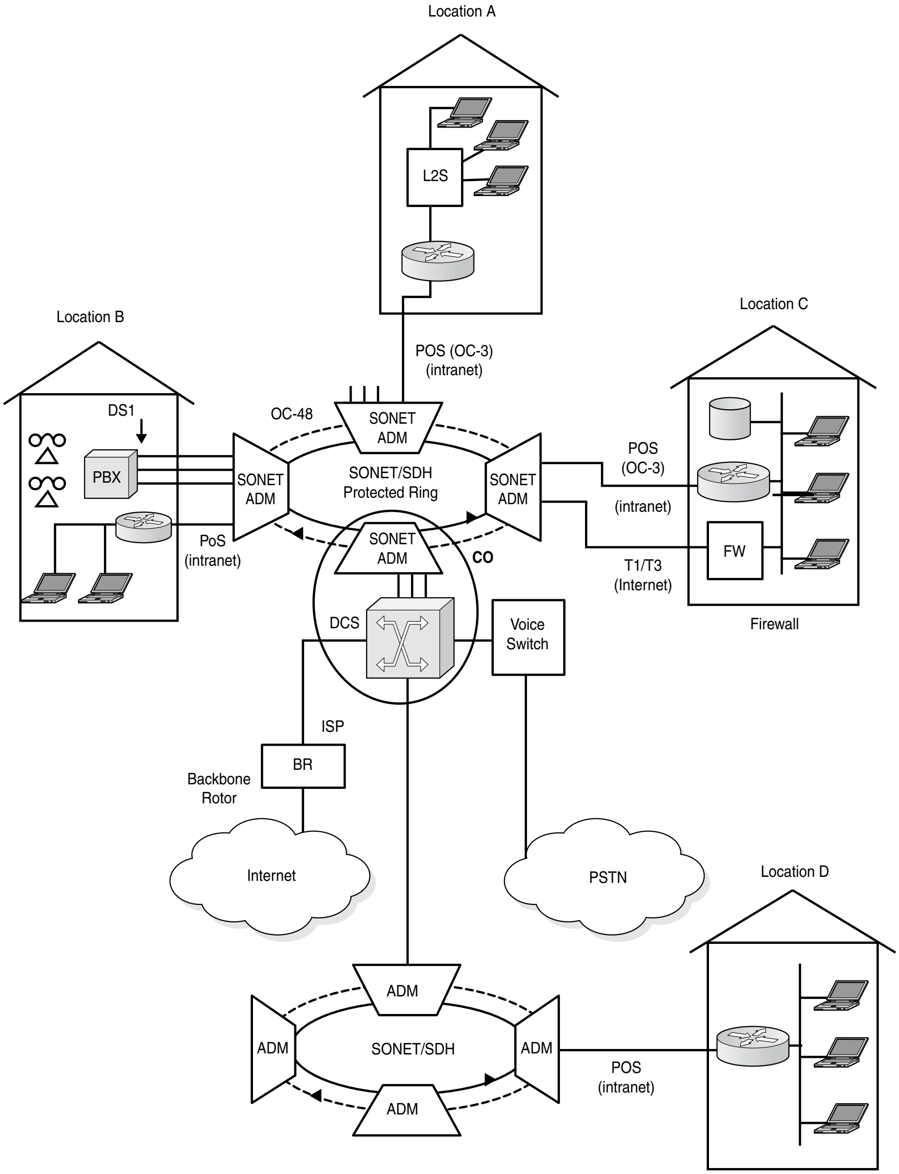


Figure 3-71
PoS encapsulation process.

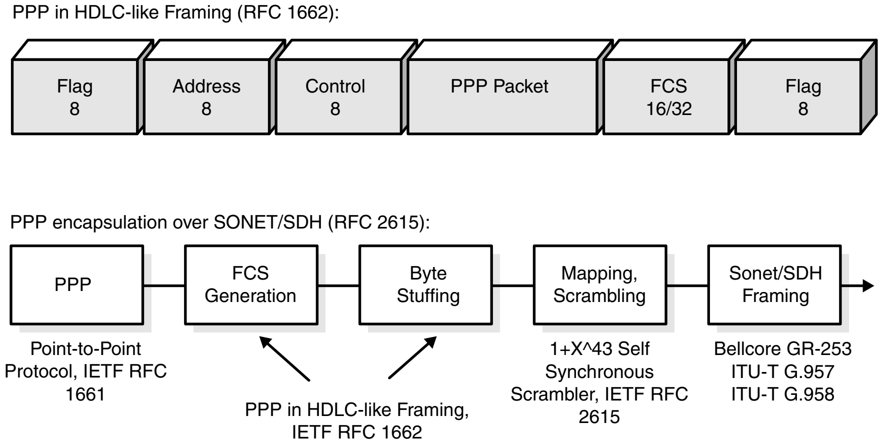


Table 3-5
IETF RFCs

		STATUS
1619	PPP over SONET/SDH	Standard protocol
1661	The Point-to-Point Protocol (PPP)	Standard protocol
1662	PPP in High-Level Data Link Control (HDLC) framing	Standard protocol
2615	PPP over SONET/SDH	Proposed standard protocol

optics). IEEE 802.17 Resilient Packet Ring (RPR) is viewed by some as a possible successor specification. For carrier networks, lambda switching is considered the next wave (see Chapter 7, "All-Optical Networks").

End Notes

1. A comprehensive discussion of SONET was provided in D. Minoli, *Enterprise Networking: Fractional T1 to SONET, Frame Relay to BISDN* (Boston: Artech House, 1993).
2. Today 10-Gbps OC-192C systems are also being deployed.
3. Stephan Koffler and Lori Franklin. "Optical Networking: A Focus on the Metro Opportunity." *First Union Securities,* New York, June 2000.

4. The excerpt that follows is from Glenn Bischoff's "Promises Made, Promises Kept," *Telephony* (July 9, 2001).

> " . . . So when ELEC 'X' decided its value proposition was going to be the provisioning of non-oversubscribed bandwidth at guaranteed transport speeds of 100 Mbps for the flat rate of $1,000 per month (including Internet connection), it was prepared for more than a little bit of skepticism. 'It is almost a too-good-to-be-true proposition," acknowledges [Va]. 'But we welcome the skepticism.' [ELEC 'X'] is very specific in terms of the companies it targets for its service: commercial end-users in multitenant office buildings, specifically those that have to transport a large number of high-capacity files such as law firms and financial services/investment companies. [ELEC 'X'] looks for buildings that have a comparatively large percentage of these types of companies. In addition, a building must have at least 100,000 square feet of office space and 30 tenants for [ELEC 'X'] to consider it a viable prospect. *However, the sweet spot is 1 million square feet of office space and 67 customers, says* [Ku]: 'We feel very confident we can sell into that building in a very profitable way.' Buildings must also be in central business districts and in close proximity to dark fiber. 'We try to determine where we can build the most cost-effective rings,' says [Ku]. To date, [ELEC 'X'] has found 540 buildings that fit the bill and has signed up more than 3,000 customers for service. Of those customers, about 60 are currently receiving service. Each city has a hub (except for New York, which has two). Each hub connects to a series of physical rings; each ring connects up to eight buildings. The rings are typically no more than 10 to 15 kilometers in circumference. To make the connections between the buildings and the hubs, [ELEC 'X'] leases metro dark fiber on an indefeasible right-of-use basis for 20 years, with a 10-year option, which [Va], [ELEC 'X']'s vice president of marketing, says gives the company all the rights of ownership and saves it money at the same time. In each building, [ELEC 'X'] sets up two OC-48 dense wave division multiplexers in a figure-eight configuration; one wavelength provides capacity of 2.5 Gbps in a clockwise direction, while the other does the same in a counterclockwise direction, without the use of optical amplifiers. The wavelengths are physically diverse and protected at Layer 3. The resulting 5 Gbps of capacity is enough to supply a maximum of 46 customers in each building with non-oversubscribed clear-channel 100 Mbps access back to the hub site. If [ELEC 'X'] signs up more than 46 customers in a building, all it has to do is drop in a few more wavelengths, says [Ku], the company's chief technology officer, who adds that [ELEC 'X']'s architecture is quite different from that used by other optical Ethernet providers. 'A more typical advanced network configuration would run a single gigabit Ethernet ring consisting of two fibers with a single Ethernet connection that

is shared among a number of buildings,' he explains. 'We are only using one strand of fiber to serve eight buildings, so we are using our capital very efficiently.' In all, it is a pretty simple deal, concludes [Ku]. 'The biggest challenge is getting the laterals into a building,' he explains. 'It is a slow process because it involves getting permits, digging up streets and dealing with regulatory issues. That is the real work in the trenches.' According to [ELEC 'X'], the privately held company has raised >$420M in funding with >$110M from the capital markets and $310M in private equity funding . . ."

Note that 60 paying customers at $12,000 a year only generate $0.7 million in annual revenue.

5. Streams such as DS1s are first bit- or byte-multiplexed into VTs.

6. Nortel Corporation, "SONET 101," White Paper, Document 56118.11, Ottawa, Canada.

7. In an asynchronous DS3 frame, the DS1s have gone through a number of levels of multiplexing that include the addition of stuffing and framing bits. The DS1 signals are mixed somewhere in the information bit fields and cannot be easily identified without completely demultiplexing the entire frame.

8. Clocking implies using a series of repetitive pulses to keep the bit rate of data constant and indicate where the ones and zeros are located in a data stream.

9. A comprehensive discussion of TMN was provided in D. Minoli et al., *Planning and Managing Corporate ATM-Based Networks,* Greenwich, CT: Prentice-Hall/Manning, 1997.

10. FOTS stands for fiber-optic transmission system.

11. SONET can be considered a subset of SDH.

12. CI is the arrangement of equipment and wiring on the customer premises (owned and maintained by the customer).

13. Nortel, "Introduction to SONET Networking—A Tutorial Handbook of Advanced SONET Networking Concepts," October 1996.

14. Nortel's S/DMS TransportNode enables up to 11 point-to-point systems to share a common protection channel (1:11 protection).

15. ATM's overhead: 5-byte tax for every 48 bytes of data sent.

CHAPTER

4

Next-Generation SONET and Optical Architectures

4.1 Motivations for Next-Generation Technology

Although traditional Synchronous Optical Network (SONET) approaches to packet data transport can handle connectivity needs in some midspeed situations (for example, 10 to 100 Mbps), they generally do not adequately address the full requirements of emerging data-centric high-capacity networks (for example, 1 to 10 Gbps). Such requirements can be better served by new technologies. Observers continue to hold the opinion that the growth of Ethernet- and IP-based packet data services will proceed unabated and that data services will soon comprise the majority of all telecommunications traffic. This shift in service mix can be attributed to an increase in the number of data connections (specifically for Internet/intranet access) and to a higher average bandwidth requirement for data applications. In the corporate landscape, these drivers translate into a rapidly accelerating demand for native Ethernet transport services over metropolitan area networks (MANs) and wide area networks (WANs). The de facto situation for many carriers is that their optical networks are generally optimized for voice transport over time-division multiplexing (TDM) channels. Although these networks can provide viable solutions in lower-end data transport applications, they are not ideal environments where packet data dominates the service mix; statistical TDM (STDM) (packet) methodologies may be better suited.

Very rarely can a single technology meet every attribute for every service. Therefore, it is important for service providers to select which technology combination best fits their business models, in both the immediate- and long-term planning cycle. High-end service providers such as InfoPort Communications Group have multiarchitecture, multitechnology, multiservice, and multi-quality of service (QoS) networks. Equipment vendors have also realized that no single technology can do it all and have begun to combine technologies into network elements that enable service providers to easily add new services.[1] Next-generation SONET/Synchronous Digital Hierarchy (SDH) systems typically go in this direction.

It is possible that next-generation SONET may enjoy easier market penetration than Ethernet-based technologies. Recently, it was observed that the competitive landscape of telecommunications in the United States has shrunk significantly with four major players that have become the center of the world.[2] The implications are that the introduction into the network of the Ethernet at the core architectures now much in the press may not be as rapid as proponents had hoped. The incumbents may feel much more comfortable with next-generation SONET platforms since these behave, by and large, like traditional SONET systems. The new players who might have used the new technologies recently have had difficulties securing venture capitalist funding. Before proceeding with an assessment of next-generation SONET, we offer our rationalizations for the point just made; this point has major implications for technology developers. Why are new players who might have made use of the new technologies having difficulties securing venture capitalist funding? We believe this is because of the following reasons:

- The Competitive Local Exchange Carriers (CLECs), Digital Subscriber Line LECs (DLECs), and Building LECs (BLECs) were carbon copies of the Incumbent Local Exchange Carriers (ILECs) in approach, architecture, and philosophy (being staffed by ex-ILEC employees and executives), without much innovation. New players need to show that they are significantly different from these CLECs, DLECs, and BLECs and afford new thinking.
- The CLECs, DLECs, and BLECs misquoted the achievable financial results, promising 90 percent gross margins to investors (doing so by inventing their own definition of gross margin). Realistically achievable metrics are 40 to 50 percent gross margins and 20 to 25 percent net returns. New players need to use the correct financial models and show that they are significantly more conservative than the CLECs, DLECs, and BLECs by underpromising and overdelivering.
- The CLECs, DLECs, and BLECs focused on narrowband services, where the chargeable fees are $20 to $50 per month per line, and were totally dependent on the ILECs for the underlying raw material, namely the copper loops, thereby greatly reducing margins and wealth-generation capabilities. To be a $600-million-a-year company (as an example), you need 1 million customers. A startup company is ill prepared to handle this volume from an operations support system (OSS) standpoint. New players need to focus on broadband services where there are more opportunities for revenues and margins and more suppliers for the raw material, namely fiber.
- The CLECs, DLECs, and BLECs took an overcapitalization approach of building much larger networks and purchasing more equipment than would have been justified by primary market research on the revenue-generation potential. Many CLECs, DLECs, and BLECs networks only had 2 to 4 percent utilization. Their architectures were not innovative and they used decade(s)-old approaches, methodologies, and designs. New players need to develop creative, lean architectures that employ

much less equipment (capital), and the players need to keep their inventory of capacity significantly lower than the 96 to 98 percent overbuilt levels that were common in the late 1990s and early 2000s.

- The venture capitalists tend to focus only on the production side of the equation, without appropriately funding the consumption side of the equation. Venture capitalists like to fund new technology development, but they have also not funded to an optimized macroeconomic level new carriers and providers that can develop markets for the new products that the venture capitalists have (over)funded on the technology side. If venture capitalists want to secure a return on their technology investments, they also need to diversify their portfolios and fund the technology usage industry. As a guideline, for every $2 invested in the technology production side, they should invest $1 in the technology consumption side. If some point of equilibrium in the funding discipline is not reached, the investments of the technology creation side will not provide returns. Naturally, these venture capitalists need to find the right startup companies to fund. As a guideline, if a carrier promises 90 percent gross margins, 30 percent market penetration, EBITDA-positive status in one year, $100 thousand in equipment for every building brought on net, $1 in annualized revenue for every $10 invested, or an unsustainable fire-sale price for services, it is tautological that these startups will not generate positive returns for venture capitalists.

4.2 Features of Next-Generation SONET

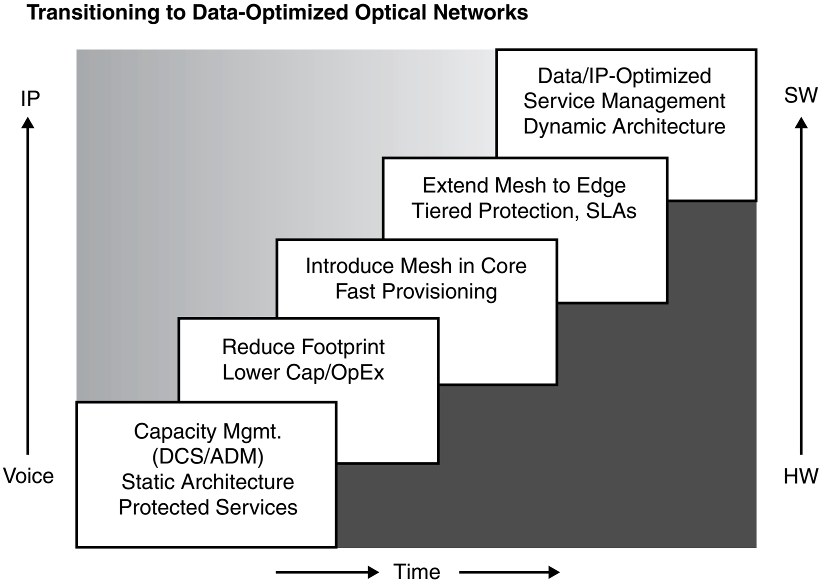
As a technology, next-generation SONET systems are not revolutionary, but rather are evolutionary. At the network topology level, these systems are generally similar or identical to traditional SONET/SDH systems. Basically, next-generation systems are data-aware SONET network elements (NEs) that enjoy one or more of the following characteristics:

- They have direct 100-Mbps, 1-Gbps, or 10-Gbps Ethernet handoffs in addition to traditional handoffs.
- Data interfaces such as ESCON, FICON, and Fibre Channel are supported.
- They are cheaper than traditional SONET NEs because of advances in technology and because they do not support voice-originated capabilities such as a complex virtual tributary (VT) mechanism.

- They support full Operations, Administration, Maintenance, and Provisioning (OAM&P) while enjoying Simple Network Management Protocol (SNMP) access to the management information.
- They support new trunk-side architectures (for example, they support a packet-based mechanism such as, but not restricted to, Resilient Packet Ring [RPR]) and have OC-x handoffs.
- They support IP routing functions, including perhaps multiprotocol label switching (MPLS).
- They obviate the need for Asynchronous Transfer Mode (ATM), but employ framing based on Packet over SONET (PoS) or the Generic Framing Procedure (GFP).

Next-generation SONET is not supported by a standardized definition by the American National Standards Institute (ANSI) or the International Telecommunications Union (ITU). Rather, it is a collection of vendor-advanced approaches to improve the cost-effectiveness of traditional SONET in a broadband data context. In 2001, vendors applied recent advances in optical components and software architectures to develop next-generation SONET. (See Figure 4-1 for a view of the evolution.)[3] For example, advances in transceivers, SONET Application-Specific Integrated Circuits (ASICs), PoS, forward error correction (FEC), digital wrappers,

Figure 4-1
Factors related to the transition to next-generation platforms.



lumped and distributed amplification, gain and dispersion equalization, optical switching, high-speed ASICs for electronic switching and grooming, and Vertical Cavity Surface-Emitting Lasers (VCSELs) were leveraged to reoptimize SONET equipment. We also saw a gradual transition from SONET rings to intelligent mesh architectures, although it was more often applicable to the long-haul and regional area network than to the metropolitan network.

Current networks' multiple layers can be difficult and costly to provision and manage. Therefore, a much simpler SONET-based approach is advantageous. Solutions include

- Optical Ethernet, namely the various virtual local area network (VLAN)/campus-extended LAN technology approach
- RPR-based systems
- Streamlined PoS
- GFP-based PoS

Multiservice solutions afforded by next-generation SONET (or by any platform for that matter) must incorporate QoS, fault tolerance, and the restoration and low latency of SONET. Additionally, it must be able to handle the bursty nature of (some) data networks while offering the scalability and capacity to transport video applications. It is also desirable for these capabilities to be added incrementally because each service provider's timing and business model may be different. Figure 4-2 is a usage example of next-generation SONET, with the implication that metro access may be where this technology makes the most sense initially.

A distinction needs to be made between an architecture that uses Ethernet at the edge (which has been deployed by carriers for over 10 years, typically using a point-to-point private line, traditional SONET, and/or ATM with distinct edge devices) and an architecture that uses Ethernet at the core (a new technology). In our definition, next-generation SONET mostly fits the former case, but it offers better streamlined data-ready/data-inclusive equipment with integrated Media Access Control (MAC) bridging/MAC tunneling and/or IP routing, and improved operational efficiency compared to traditional SONET systems. Our definition, however, also includes the reverse case, where the NE is packet based at the core, but it has both Ethernet as well as traditional ITU SONET handoffs at the edge. In the first case, a packet path is emulated over a circuit path; in the latter case, a circuit path is emulated over a packet path.[4] Table 4-1 depicts the various combinations of what we call next-generation SONET for the purpose of this investigation. As noted, there is no standard definition (that is,

Figure 4-2
Application example for next-generation SONET.

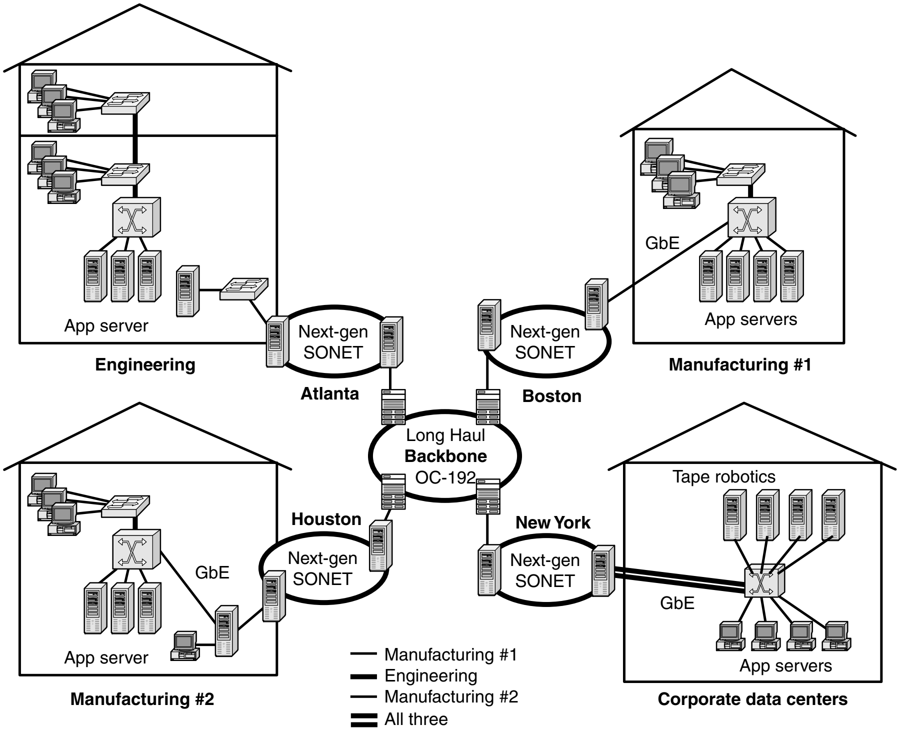


Table 4-1

Various Topo-architectural Combinations

Edge (User Handoff)	Core (Network Handoff)	
SONET	SONET	Traditional SONET[(a)]
Ethernet	SONET	
SONET Ethernet	Packet (RPR)	Next-generation SONET
SONET Ethernet	Hybrid: packet (RPR) on some slots and TDM on other slots	
SONET Ethernet	Ethernet	Ethernet based[(b)]
Ethernet	Ethernet	

[(a)] Discussed in Chapter 3, "Traditional SONET Architectures"

[(b)] Discussed in the companion book

published by a standards body) for “next generation.” Various vendors have developed many packaging methods for functions deemed to be of value to data-intensive metro access/metro core applications.

According to market research, optical Ethernet services in the MAN are expected to take an increasing share of the lit services market, rising to 20 percent by 2005 and representing approximately $6 billion in revenues.[5] This implies that the total metro optical market will be $30 billion in 2005. Pioneer Consulting predicts that the next-generation SONET market will grow to approximately $6 billion by 2004 (representing another 20 percent of that total; the implication would then be that the balance, or 60 percent, remains in traditional SONET equipment).[1] The reasons for the growth in next-generation SONET are quoted as follows:

- Capability to interoperate with existing SONET networks
- New system architectures that enable simple capacity and service upgrades
- New enabling technologies that increase network revenue-generating capacity by as much as an order of magnitude
- Native data and video interfaces that reduce the number of NEs
- Functionality that combines digital crossconnect systems (DCS), edge routing, and add/drop multiplexing (ADM) functionality into a single NE for reduced overlay and purpose-built networks

Figure 4-3 is a protocol-model view of both traditional and next-generation SONET. Figure 4-4 illustrates what a typical next-generation SONET network element might look like from a plug-in standpoint. Typical functions include time slot assignment (TSA), time slot interchange (TSI), hairpinning, broadcast, and drop and continue; support of packet-mode trunking is also included in some products.

In next-generation SONET, you look for cheaper transport and network elements, but with a full complement of OAM&P capabilities. Because of its emphasis on customer-ready interfaces such as Ethernet, next-generation SONET is generally more applicable to the metro access portion of the network than the metro core (see Figure 4-5).

Examples of next-generation SONET equipment available in early 2002 included the Cisco ONS 15000 family and the Nortel Networks’ OPTera Metro family. Figures 4-6 and 4-7 show the typical internal architecture for a next-generation SONET system—the former employing packet technology on the trunk side and the latter employing a traditional TDM technology.

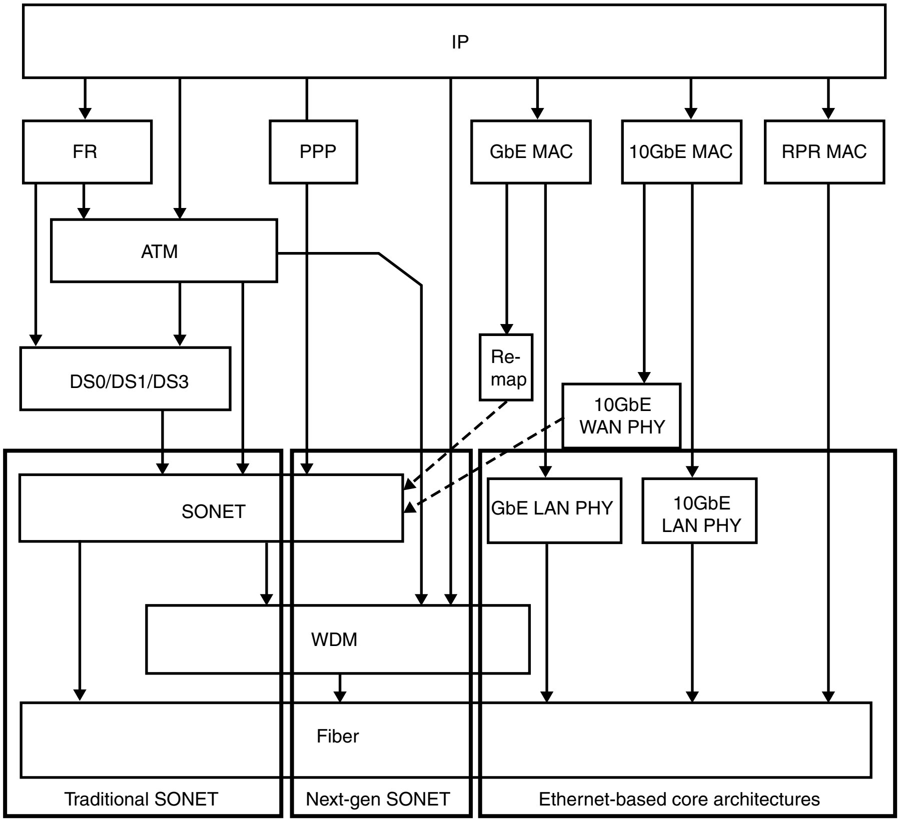


Figure 4-3 Data-aware SONET/SDH.

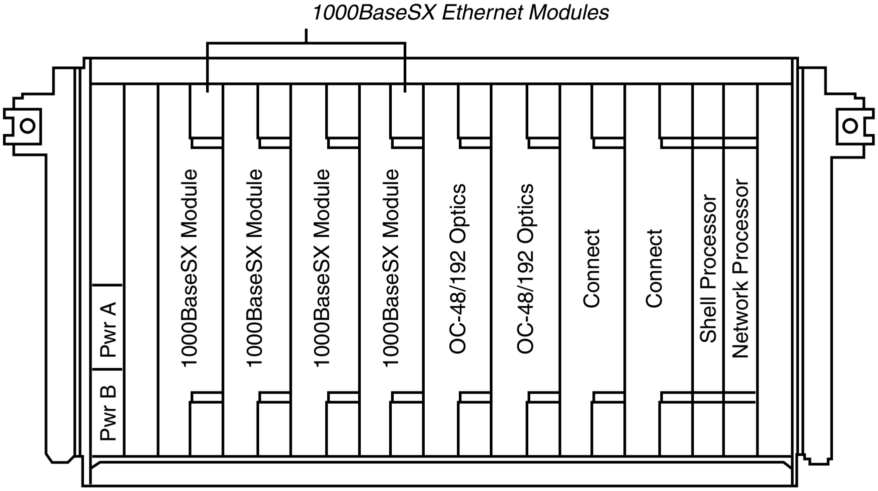
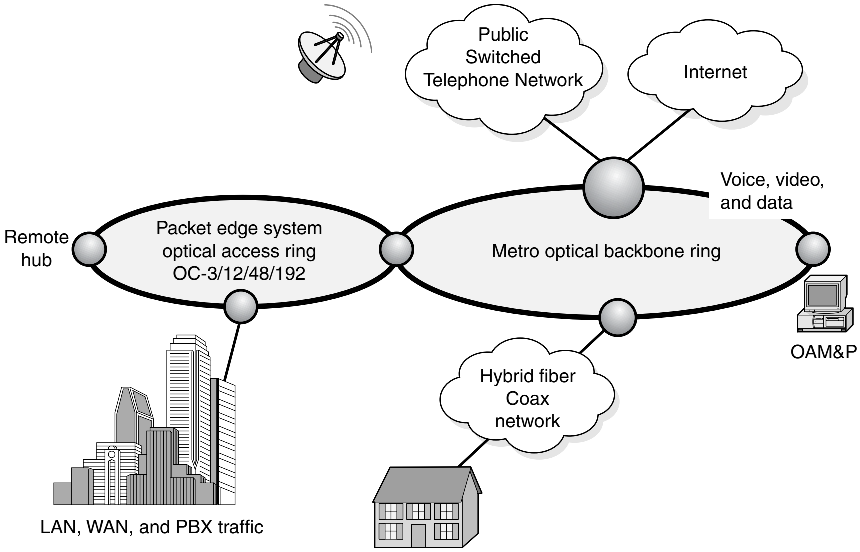


Figure 4-4 Typical plug-in array for a next-generation SONET/SDH.

Figure 4-5 Typical topology for next-generation SONET.

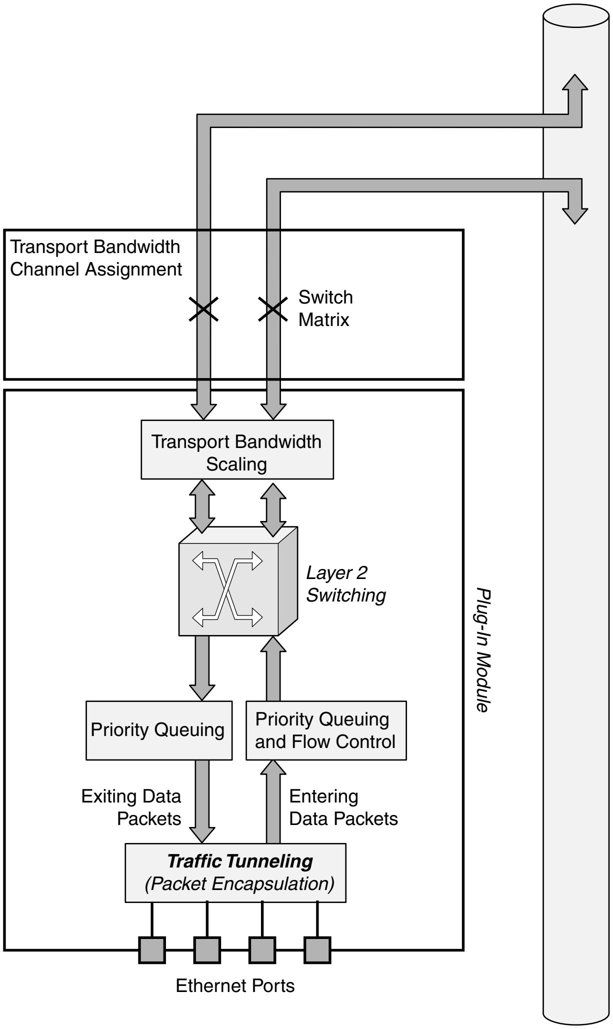


4.3 Financial Implications

Proponents make the case that in the face of new startup providers offering broadband services, incumbent carriers must complement their private-line offerings with Ethernet over SONET (EoS) services. Proponents say that these incumbents "have the edge over their new rivals in that they own metro SONET rings and long-haul backbones, and have deep pockets."[6] Thus, an overlay of edge rings that utilize next-generation SONET may be useful. However, the new startup providers have simpler networks to manage and a chance to prove that they can quickly provision, maintain, and service optical/Ethernet connections while turning a profit.[6]

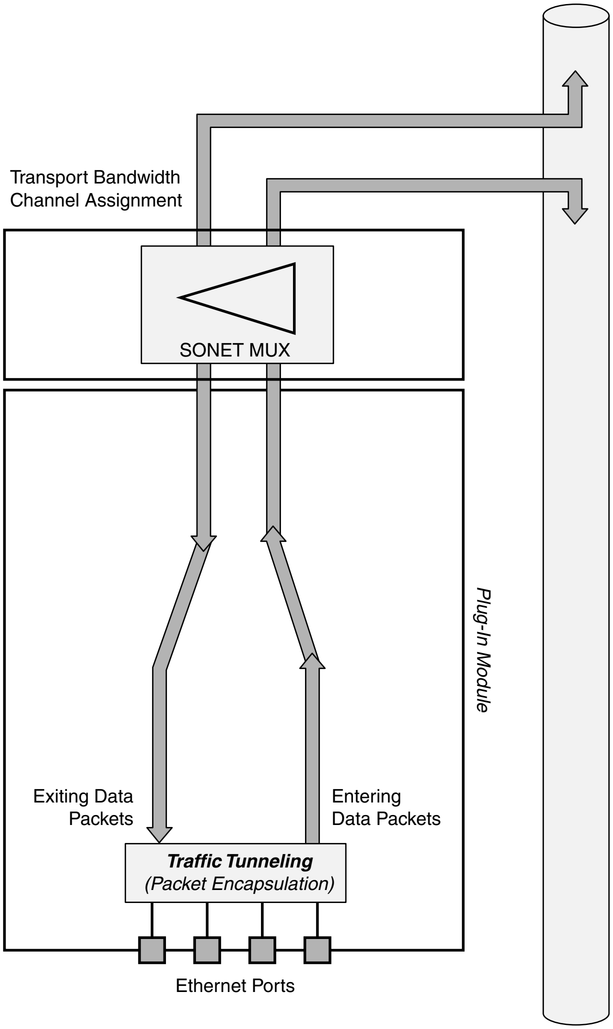
Critical to an understanding on the part of the equipment manufacturers as to what kind of equipment will be purchased by service providers is the concept we introduced in Chapter 1, "Advances and Opportunities in Next-Generation SONET and Other Optical Architectures." In Chapter 1, we invoked the principle of "following the money." In rendering an end-to-end reliable secure monitored service, where are the cost components that comprise the total service cost? Is the cost in the protocol engine or is it in the physical fiber link from the building to the metro core network? Figure 4-8 provides a hint and subtext for equipment manufacturers. The figure implies that a mere cost reduction in the edge network element (even at a

Figure 4-6
Next-generation SONET system—STDM trunk.



100 percent rate), either because the price of the SONET NE is reduced or because it eliminates a supplemental bridge/router at the customer building, only has a 6 percent impact on the total monthly recurring charge (MRC) cost to provide transparent LAN service (TLS) to 4 customers in a building. On the other hand, if the new technology can reduce the operations cost by 100 percent, then the overall reduction in service cost is about 10 percent. If, optimally, the equipment cost is reduced by 100 percent and the operations cost is reduced by 100 percent, then the overall cost of the

Figure 4-7
Next-generation SONET system—TDM trunk.



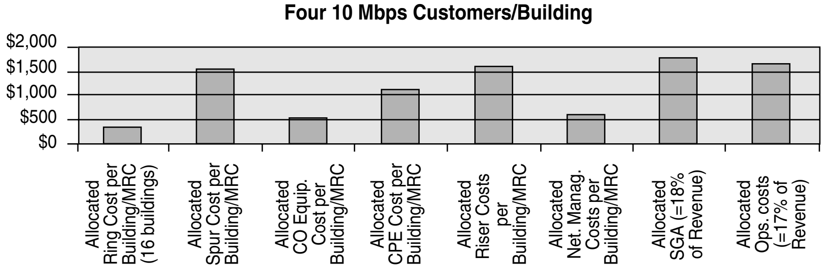
service is reduced by 15 percent. This last figure of merit, if reflected in the price of the broadband service quoted to the customer, is closer to the ballpark of where an end user might make the decision to switch providers. It should be noted that the Ethernet LECs (ELECs) are targeting midsize to large business users for TLS using VLAN-based campus-extended Gigabit Ethernet (GbE). However, most service providers need plans to offer the "triple play" of voice, data, and video on demand if they are to remain viable as businesses. Hence, it is important to have a technology that serves all of

Figure 4-8 *Advantages of edge NE cost reduction versus operations cost reduction.*

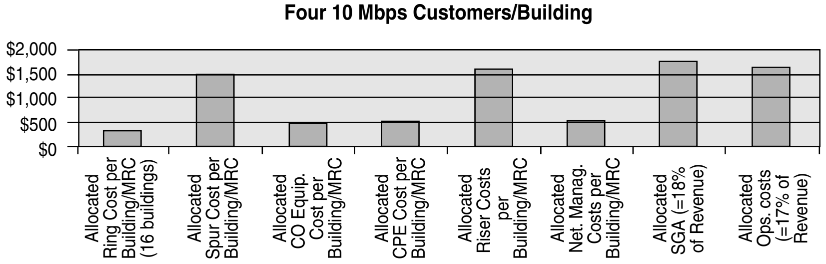
A = Total cost to support four broadband users in a building, with edge NE cost at $40,000: $9,124
B = Total cost with edge NE cost at $20,000: $8,568
C = Total cost with edge NE cost at $40,000 and operations cost reduced to 50% of above: $8,274
D = Total cost with edge NE cost at $20,000 and operations cost reduced to 50% of above: $7,718

(A) (B): –6.0%
(A) (C): –9.3%
(A) (D): –15.4%

Total cost to support four broadband users in a building, with CPE cost at $40,000: $9,124



Total cost with CPE cost at $20,000: $8,568

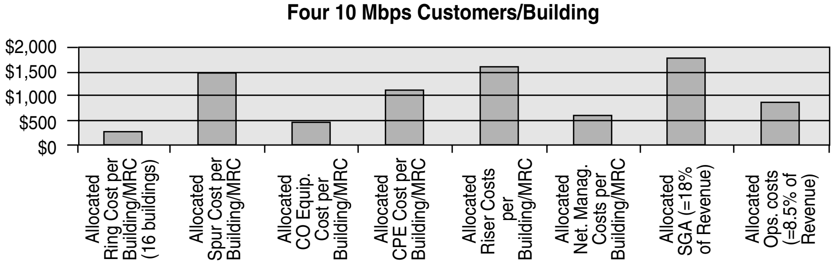


these requirements. Next-generation SONET provides a viable option to serve all of these requirements.

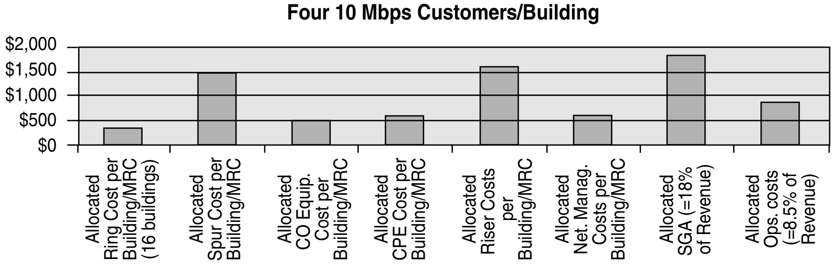
For carriers to justify connecting a building to the metro core network, they need a viable business case. A macro-level model of the viability of a sustainable ELEC (that is, a non-year-by-year venture-capitalist-fed ELEC registering hundreds of million-dollar losses per year) is simply as follows. Assuming a $100,000 fiber-installation investment to connect a building, a $50,000 in-building edge multiplexer investment, a 30-month payback target,[7] and three broadband customers in a building that the specific carrier in question can bring onboard, the minimum per-month charge per customer would have to be $1,666. This figure does not include the metro core

Figure 4-8 cont. *Advantages of edge NE cost reduction versus operations cost reduction.*

Total cost with CPE cost at $40,000 and operations cost reduced to 50% of above: $8,274



Total cost with CPE cost at $20,000 and operations cost redu ced to 50% of above: $7,718



network, the central office/service node/points of presence (CO/SN/POP) equipment cost, the long-haul/Internet cost, the operations cost, and the sales, general, and administrative (SGA) cost. On top of this, you need to add some amount of net profit for the carrier. Thus, a figure of $5,000 MRC or higher per broadband customer may be the minimum sustainable figure that a carrier can charge over the long term.[8] Hence, there is a high degree of disincentive on the part of carriers to offer a billing arrangement that differentiates bandwidth delivery down to the 1-Mbps level, and frankly, few user applications are sensitive to a 1-Mbps granularity. In turn, this translates into a disincentive to wire a building with fiber and, consequently, into a general disincentive to deliver broadband services. Sure, if a carrier was to offer an MRC with the formula $4,800 + 25×m, where *m* is the number of megabits that the customer wanted, this could be workable from the carrier's point of view; however, the granularity may be sufficient to the levels discussed previously and not at the 1-Mbps level.

Incumbent providers that also support traditional lower-speed data services beyond TLS have additional design considerations. Traditional approaches generally do not adequately address the broad mix of traditional circuit-based services. For incumbents considering PoS channels, traditional ATM, or Frame Relay, serious drawbacks in the critical areas of cost-effectiveness, scalability, manageability, survivability, and service flexibility quickly become apparent when compared to next-generation SONET solutions, as shown in Table 4-2. Next-generation systems provide an improved IP delivery solution. According to Nortel Networks' statistics, these systems can provide savings to the operator of up to 70 percent over

Table 4-2

Comparing Architectural Alternatives

Packet Transport Alternatives	**Drawbacks/*Advantages***
Traditional ATM	■ More expensive ■ No built-in redundancy ■ Many hard-to-manage permanent virtual circuits (PVCs) required
Frame Relay	■ Limited bandwidth ■ Many hard-to-manage PVCs required
PoS channels	■ Not scalable ■ Inefficient bandwidth usage ■ Many hard-to-manage point-to-point connections required
Fiber-Optic Interrepeater Link (FOIRL)	■ Limited bandwidth ■ Primarily short-haul solution ■ Inefficient use of fiber resources ■ No built-in redundancy ■ No end-to-end management
Router-based packet transport rings	■ Limited support for mixed services ■ Limited support for long-haul, high-capacity applications ■ Not based on carrier-grade network elements ■ Proprietary technology
Next-generation SONET	■ *Scalable* ■ *Cost-effective, bandwidth-efficient* ■ *Connectionless* ■ *Full support for mixed packet data and TSM services* ■ *Short-haul or long-haul, high capacity up to 1.6 Tbps* ■ *Survivable with inherent resilience* ■ *Easily manageable* ■ *Carrier-grade equipment* ■ *Open industry standards*

traditional solutions.[9] Multiservice next-generation SONET platforms enable the end-to-end delivery of multiple high-speed services, including 10 GbE, PoS, Fibre Channel, ESCON, FICON, and GbE. These platforms enable the delivery of traditional and new service offerings such as data center interconnections, storage area networking (SAN), VLANs, and virtual private networks (VPNs). The new platforms provide low-cost native LAN interfaces, eliminating the need for more expensive TDM transport interfaces in customer equipment and/or additional customer premises systems (such as T1/T3 multiplexers) in order to access the backbone. Typical equipment available at the time of this writing possesses the following features:

- Carrier-grade reliability
- Protocol and bit-rate independence
- Ethernet switching integrated directly into the optical layer
- Up to 32 protected wavelengths and 64 unprotected wavelengths
- Per-wavelength protection switching
- Dense wave division multiplexing (DWDM) bandwidth scalable to 2.5 Gbps per wavelength (10 Gbps in the future)

The ELECs startup providers typically aggregate Ethernet traffic from customers with IP routers and Layer 3 switches, and then direct the traffic to a core router in the packet-based long-haul carrier's point of presence (POP). Proponents observe that the new startup providers are at the mercy of the carriers' private backbones in offering end-to-end reliability and QoS, but can gain more control by placing GbE switches that encapsulate Ethernet packets in DWDM wavelengths in the backbone carriers' POPs. This Ethernet over DWDM implementation addresses reliability and QoS issues, but it comes with a higher price that customers may not be willing to pay.[6] The use of next-generation SONET, on the other hand, can address many of these QoS and traffic-engineering issues. Again, as a backdrop, it is important to keep in mind that customers are not willing to migrate from a $300-a-month Frame Relay connection to a $3,000-a-month broadband connection if the quality, reliability, security, and availability leaves something to be desired.

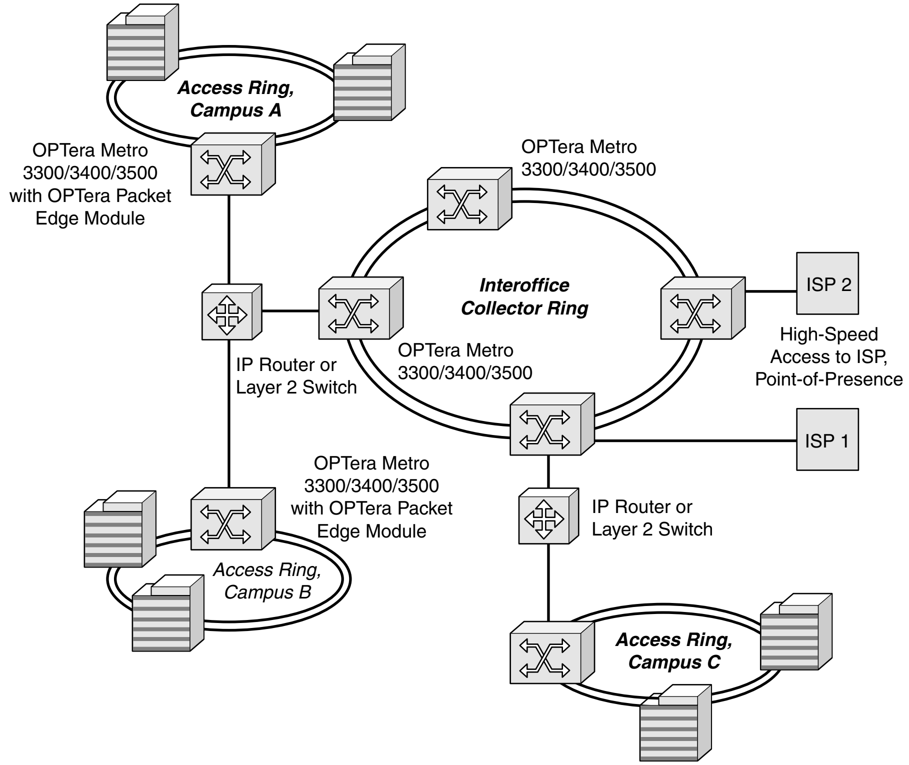
Manufacturers of equipment are publishing white papers cheerleading or sable-rattling the ILECs with statements such as these: "Though the bulk of venture capital funding has declined . . . optical networks continue to attract strong equity financing. Several metro providers utilizing optical Ethernet-based solutions have secured significant funding in 2001. Not only are these emerging carriers financed, they are focused, aggressive, and operating in the top 20 U.S. metropolitan markets. . . . It is recommended that the ILEC introduce optical Ethernet services initially at 100 Mbps and above to minimize the potential impacts of cannibalization on some existing data services such as T1 and T3. Service-offering and pricing strategies enable the introduction of optical Ethernet while minimizing impact on the current revenue base. Optical Ethernet services complement the ILEC service set by offering new revenue opportunities through the ability to offer scalable customer-tunable, high-speed data services, as well as a cost-effective platform to deliver new value-added services, such as Voice over IP, Network Attached Storage, and Internet Data Centers . . . These emerging Ethernet carriers have used their facilities-based CLEC models, along with optical Ethernet technology, to carve a path into the RBOC's stranglehold on metro access . . ." Equipment manufacturers are quoting an internal rate of return (IRR) of 87.6 percent and a payback in less than 18 months. This kind of advocacy is intended to get ILECs to purchase the new next-generation SONET or other packet-mode network elements. However, the cited financial metrics are unachievable in our opinion when the total cost of rendering service is appropriately taken into account, such as fiber riser costs, building-owner costs, building fiber spur costs, metro backbone fiber costs, building and CO equipment costs, CO rent/power costs, network management costs, engineering and provisioning costs, other operations costs, and SGA costs.

4.4 Approaches

In the past, carriers typically built overlay networks to accommodate different characteristics of various media (voice, data, video, and the Internet) because there was no single technology or product that could meet the requirements for each service. New SONET systems are aimed at addressing these requirements. In addition, they can be used in conjunction with embedded systems, as Figure 4-9 demonstrates.

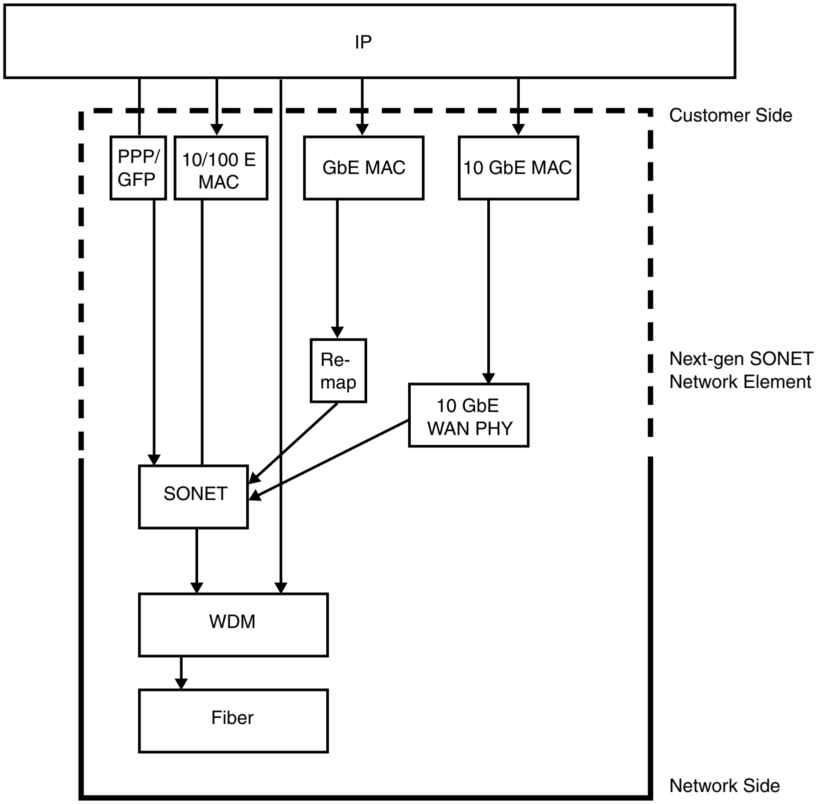
As Figure 4-10 indicates, there are a number of variants of next-generation SONET products. Traditional SONET provides a robust physical-layer protection, but it employs a rigid multiplexing hierarchy that is not well suited to data-centric line rates (although it has definitions for VTs for DS1s and DS2s, as briefly discussed in Chapter 3). For example, a 10Base-T Ethernet interface (10 Mbps) delivered in a SONET network is traditionally mapped to an STS-1 (51 Mbps), effectively wasting 40 Mbps of

Figure 4-9 Use of next-generation SONET with an embedded SONET metro core ring.



link capacity, or 80 percent. The 100-Mbps Ethernet is mapped to an STS-3c (155 Mbps). However, savvy planners understand that as bandwidth becomes a commodity, the engineering urge to optimize just one variable (namely, bandwidth efficiency) is counterproductive in the total scheme of things, much the same way PC suppliers have come to understand that it is better to be somewhat memory-usage inefficient with PCs (having 64MB for RAM or more) in favor of user simplicity via the use of a GUI-based OS such as Windows.[10] One approach to addressing trunk-side bandwidth efficiency is to employ STDM rather than TDM means. This could save, say, 50 percent of the metro core bandwidth. However, a typical metro-level backbone ring costs $5,000 MRC and can easily support 16 buildings; this equates to $313 per building per month. Improving efficiency by 100 percent would reduce the backbone cost to $156 per month. Although we are not implying that you should ignore savings here and there, the per-building cost to support 4 broadband TLS customers was in the $9,000 per

Figure 4-10 Variants of next-generation SONET.



month range; hence, a $156 reduction represents about a 1 percent overall savings. This implies that the real value of next-generation SONET systems must be in providing new services, followed by reducing operations costs, NE costs, and backbone transmission costs (through better bandwidth efficiency). Pure bandwidth efficiency by itself is of limited value.

Edge solutions need to support flexible 10-Mbps, 100-Mbps, and 1-Gbps interfaces in order to meet a wide range of end-user requirements. Occasionally, user information rates may need to be assigned in increments of 10 Mbps for flexible service level agreements (SLAs) and additional scalability. However, as bandwidth becomes a commodity, the precise (perhaps rationed) amount of bandwidth on a link becomes less of a differentiating factor for various providers. Value-added solutions will be of increased interest to the end user compared to just pure bandwidth solutions. Frankly, there is relatively little incentive for a provider that has just incurred a nontrivial expense bringing a building online to fine-tune the billing in order to distinguish, for example, between a customer's needs requiring 58 Mbps today, 59 Mbps tomorrow, and 57 Mbps the day after tomorrow. It appears at the practitioner's level that customers are more likely to be interested in 10 Mbps on a 10-Mbps interface; 50 Mbps or 100 Mbps on a 100-Mbps interface; and 50 Mbps, 100 Mbps, 200 Mbps, 500 Mbps, or 1,000 Mbps on a 1,000-Mbps interface. The fine-grain 1-Mbps granularity that some equipment vendors are bringing out appears to have rather limited commercial value. After all, if bandwidth is expected to be so abundant in the future, why be so precise about the amount of the commodity delivered?

Next-generation SONET technologies leverage virtual concatenation in order to right-size the bandwidth of the data-centric connection. We discussed in Chapter 2, "Optics Primer," the existence of a mechanism (VT) to map DS1 and DS2 signals onto the SONET payload in an efficient manner. Some people argue that there is a further need to define VTs that are data oriented—for example, 10 and 100 Mbps (and combinations thereof). For example, a 10Base-T interface could be mapped to 7 SONET VT1.5s, resulting in an increase of approximately 95 percent in transport efficiency. Vendors leveraging virtual concatenation typically provide a software-based approach to the creation, deletion, and modification of these right-sized SONET connections. Such technology enables new services such as user-based subscription and bandwidth on demand.[1] Although bandwidth efficiency is important, we have noted several times in this book that transport bandwidth efficiency is not the linchpin that will make or break the deployment of broadband to the enterprise level. Enterprise locations primarily need affordable, timely, reliable, broadband connectivity; efficiency of the payload is just an element of the overall desiderata. ATM was the multiplexing (read-efficient payload) par excellence. Once the cell tax was paid, near-perfect packing of any bursty stream can be achieved, typically with overall overbooking ratios of 5 or 10 (which more than compensate for the 20 percent or so of tax overhead). Yet ATM did not see major deployment just on the basis of bandwidth efficiency and QoS.[11] The main reason was the operational cost for both the user and the carrier; both the user and the carrier had to deploy completely new equipment and establish new procedures (at both ends) for fault management, capacity management, accounting management, performance management, and security management. The overall operational cost proved to be too high in spite of the major bandwidth efficiency achieved with overbooking ratios of 5 to 10 times.

4.5 Implementation Example

As noted previously, next-generation SONET products take one of three forms, as shown in Table 4-3.

In this section, we look at one possible implementation philosophy for a hybrid solution. The following list briefly describes an example of a next-generation SONET product family (from the Nortel Networks application brief, "Optical Internet Cable Solutions") and Figure 4-11 describes topological deployment.

- **Nortel OPTera Metro 3100** Compact for small business, this offers an OC-3 line rate with available tributaries, including 8 DS1s plus 12 DS1s, 1 DS3/EC1, or 1 OC-3/3c.

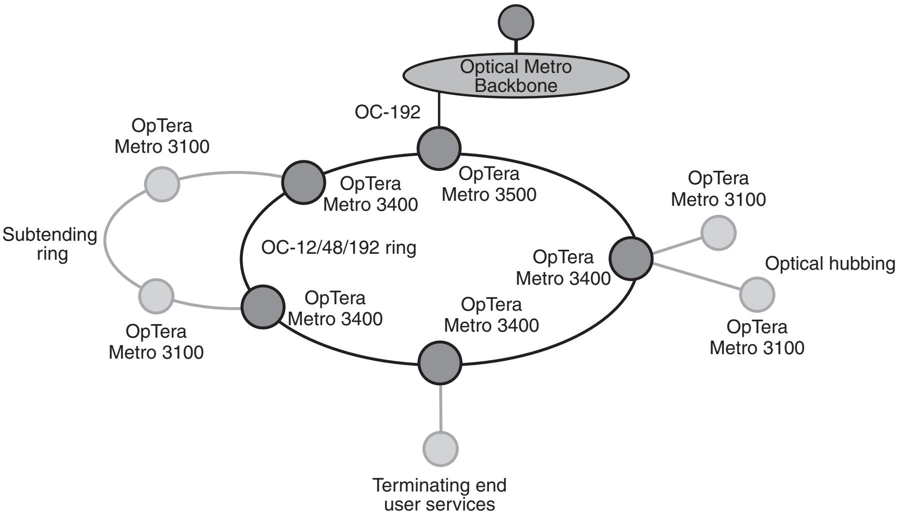


Figure 4-11 Typical topological deployment of next-generation product line.

Table 4-3 Three Forms of Next-Generation SONET Products

Edge (User Handoff)	Core (Network Handoff)	
Ethernet	SONET	
SONET Ethernet	Packet (RPR)	Next-generation SONET
SONET Ethernet	Hybrid: packet (RPR) on some slots and TDM other slots	

- **Nortel OPTera Metro 3400** Targeted for medium-sized businesses, this offers an OC3/12 line rate with available tributaries, including up to 84 DS1s, 4 DS3s/EC1s, 4 OC-3s/3cs, 8 100Base-T Ethernet, or mixed.
- **Nortel OPTera Metro 3500** This can be used for head-end applications with a high-density drop requirement. It offers an OC3/12/48 line rate with an in-service upgrade plan to OC-192. Sixteen λ DWDM capability ensures future bandwidth scalability. Available tributaries include up to 84 DS1s, 48 DS3s/ECs, 32 OC-3s, 8 OC-12s, 32 100Base-Ts, 8 GbE, or mixed (utilizes RPR).

A set of distributed packet switch modules integrated into the Nortel OPTera Metro 3000 Multiservice Platform series enables connectionless, packet-based, and shared-bandwidth networking.

With this approach to improved metro access/metro core support, packet-based data-networking modules can be added to a SONET platform to enable it to support a packetized/STDM uplink on the network side. These plug-in modules transform Nortel OPTera Metro 3300/3400/3500 Multiservice Platform and TransportNode OC-48 classic network elements into hybrid systems capable of concurrently supporting high-speed data services, voice traffic, and other traditional TDM (circuit-based) service offerings. With these plug-ins, carriers and service providers can deploy and upgrade a single network infrastructure that delivers a wide range of revenue-generating, resource-efficient data and voice services to their customers.[12] Although data packets can always be mapped into traditional leased- or private-line circuits, these solutions are often costly and inefficient from both the carrier and customer's point of view. Poor bandwidth utilization, lack of scalability, large numbers of hard-to-manage dedicated connections in mesh networks, and limited end-to-end management strategies are just a few of the many drawbacks.

With these hybrid systems, transport bandwidth is provisionable from 50 Mbps up to 10 Gbps, enabling a carrier to add more users or increase subscriber bandwidth without reengineering the network. Traffic from different locations or customers along the ring is statistically multiplexed onto a shared pool of transport bandwidth, thus achieving bandwidth usage efficiency. These systems are pre-RPR (IEEE 802.17) technologies, but incorporate many of the principles and ideas of the RPR philosophy. To ensure service continuity in the event of a network fault, Layer 2 automatic protection switching (APS) reroutes traffic, typically in about 50 ms. A self-healing ring architecture provides the geographic diversity needed to ensure the survivability of critical data services. A connectionless packet data backbone ring with distributed Layer 2 switching inherently supports the full/partial mesh traffic patterns typical of MAN/WAN applications. There is no need to provision and maintain large numbers of dedicated point-to-point connections on the backbone for mesh traffic. All fiber pairs and transport channels are available for data traffic without the requirement to reserve capacity for protection (as in circuit-based transport applications); this provides for backbone bandwidth doubling.

A typical system may support an OC3/12 line rate with available tributaries including up to 84 DS1s, 4 DS3s/EC1s, 4 OC-3s/3cs, 8 100Base-T Ethernet, or mixed. Some of the benefits might include

- A 50 to 70 percent reduction in capital cost from present methods of operation due to the elimination of unnecessary routers and the utilization of lower-cost native Ethernet interfaces
- 10/100Base-T, 100Base-FX, and GbE interfaces
- Cost-effective data networking with carrier-grade reliability
- Capability for multiple packet edge streams per SONET ring to assign a dedicated STS3/6/9/12c stream to a hub site
- Capability to group multiple hub sites to STS-n stream

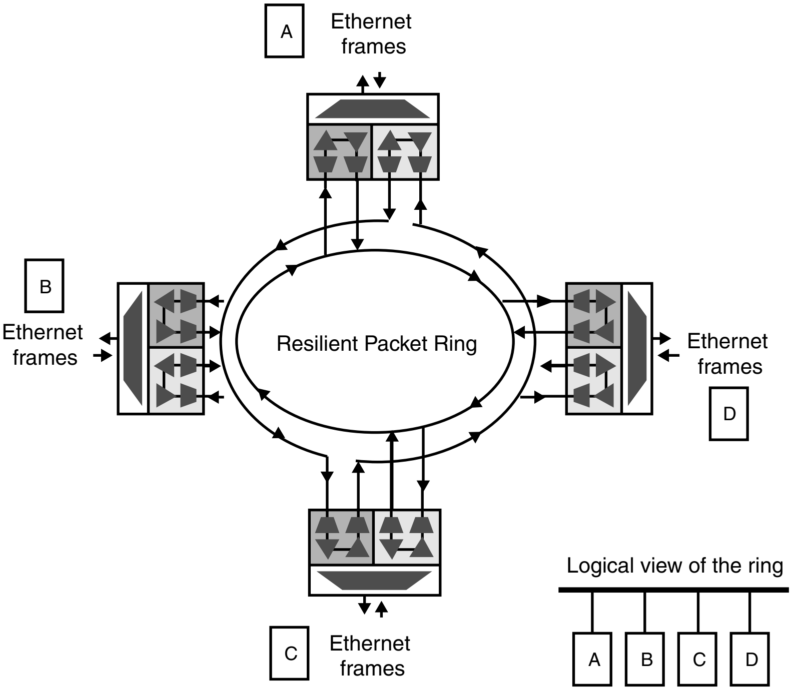
As an elaboration of this approach, you could develop equipment that was fully hybrid, where, for example, the backbone ran a SONET OC-192c. A number of constituent subchannel OC-48cs (for example, three of them) could be utilized to support crossconnected user-side OC-x to the network-side OC-48s running on the OC-192 bearer with a TDM-to-TDM mapping, whereas other subchannel OC-48s (for example, one of them) could be utilized to support statistically aggregated user-side Ethernet interfaces to the network-side OC-48 on the OC-192 bearer with a Ethernet-to-packet (for example, RPR) mapping.

Nortel Networks' OPTera Packet Edge combines three complementary, next-generation technologies:

- *Self-healing IEEE 802.17 RPRs with distributed Layer 2 switching* RPR converts a portion of a fiber ring (in STS-1 size chunks) into a logical Ethernet segment. The remaining bandwidth, if any, can be used for traditional circuit-oriented traffic (see Figure 4-12).

  This feature offers a self-healing topology that protects against fiber cuts and node failures by providing duplicate, geographically diverse paths for packet data transport. If a cable cut or other fault impairs a preferred primary route, Layer 2 protection switching automatically forwards the affected data packets to the alternate route on the opposite side of the ring. Switching to the protection route typically

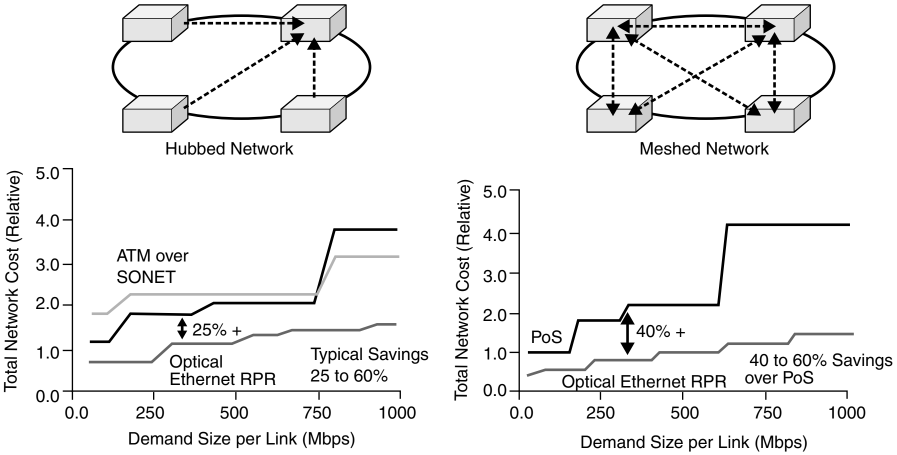
Figure 4-12
RPR environment.



requires about 50 ms so no noticeable interruption in service occurs. The self-healing RPR architecture does not require setting aside half of its fiber-span capacity for protection as in traditional SONET rings that transport voice and other circuit-based services. This yields an effective doubling of the bandwidth available for working traffic relative to circuit-based service transport. Mixed packet data and TDM circuit-based services may be placed on the same ring as desired. In these applications, circuit-based services are protected at Layer 1 using standard SONET unidirectional path-switched rings (UPSR) or bidirectional line-switched rings (BLSR) protection switching. Thus, both Layer 1 and Layer 2 protection switching can coexist on the same ring as needed to concurrently support a broad mix of service types. A more inclusive treatment is available in the book *Ethernet-Based Metro Area Networks* by Minoli and Johnson.[13]

Figure 4-13 plots savings as claimed by the vendor.[14] Naturally, any savings need to be passed onto the customer if he or she is to become interested—a 20 to 25 percent decrease in price (based on a 20 to 25 percent decrease in cost shown in the figure) is just about the savings

Figure 4-13
RPR savings claimed by vendors.



Packet over SONET (PoS), where routers interface to the transport system via TDM interfaces (DS1 to OC-12 and beyond)

ATM over SONET, where ATM switches aggregate traffic from routers and interfaces with the transport system via TDM interfaces (DS1 to OC-12 and beyond)

Optical Ethernet with RPR, where routers interface to the transport system via Ethernet, fast Ethernet, or Gigabit Ethernet interfaces on RPR NEs

level that could motivate an enterprise customer to consider switching his or her service from an existing provider/service to another. It's unclear under what circumstances, if any, an existing provider would lower its own price and cannibalize an existing PoS-delivered circuit to an RPR-based circuit.

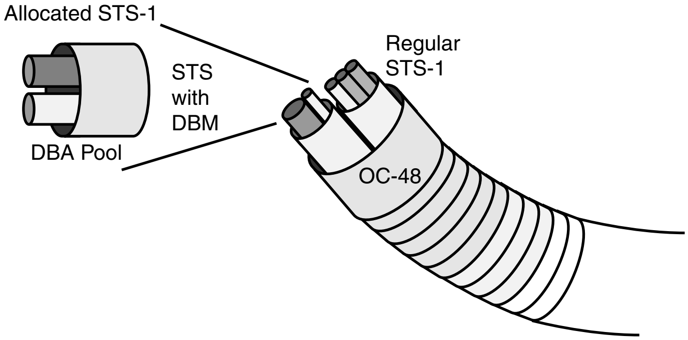
- *Traffic tunneling for efficient, secure sharing of transport bandwidth* Traffic-tunneling technology enables data from multiple enterprise customers to efficiently share common transport bandwidth on the ring while maintaining security for each customer's data packets. Also, traffic tunneling maintains traffic segregation even when customers use overlapping addresses. Traffic tunneling entails the encapsulation of entering customer data packets with customer/ring-specific address information. The added addressing identifies the correct exit point on the ring (node and port) for each entering packet. The customer/ring-specific information is automatically removed at the destination. Routing along the ring is based solely on the added customer/ring-specific address information so each customer receives only the data packets that belong to that enterprise. The tunneling process is self-learning in that packets reach their intended destinations without the labor-intensive provisioning of end-to-end connections.

- *Intelligent flow control to regulate traffic entry and enable class of service (CoS)/QoS support* Extensive flow controls regulate traffic entry and ensure that the delivered information rate conforms to the subscriber's SLA. These mechanisms also prevent certain nodes or traffic sources from unintentionally taking precedence over and starving others. The flow control mechanisms fall into three categories: the token-bucket traffic entry model, the buffer insertion ring, and the ring access protocol.

4.6 Other Approaches

Some vendors include a hardware-resident distributed algorithm that executes a distributed dynamic bandwidth allocation (DBA) protocol in the overhead of standard SONET frames. See Figure 4-14. This approach goes in the direction of RPR, but is not the exact same concept. According to proponents, this method provides a scalable, adaptive, and granular approach to managing bandwidth in a near mathematically optical fashion on SONET access rings.[1] The adaptive nature of this approach, which is managed at the physical layer, circumvents the need for complex traffic engineering on access rings. According to these same proponents, the service provider's SONET access rings leveraging this mechanism can be considered future-proofed against new and unforeseen application traffic that may adversely impact network loading. Finally, this real-time approach to virtual concatenation not only achieves new levels of efficiency in access-

Figure 4-14
DBA approach.

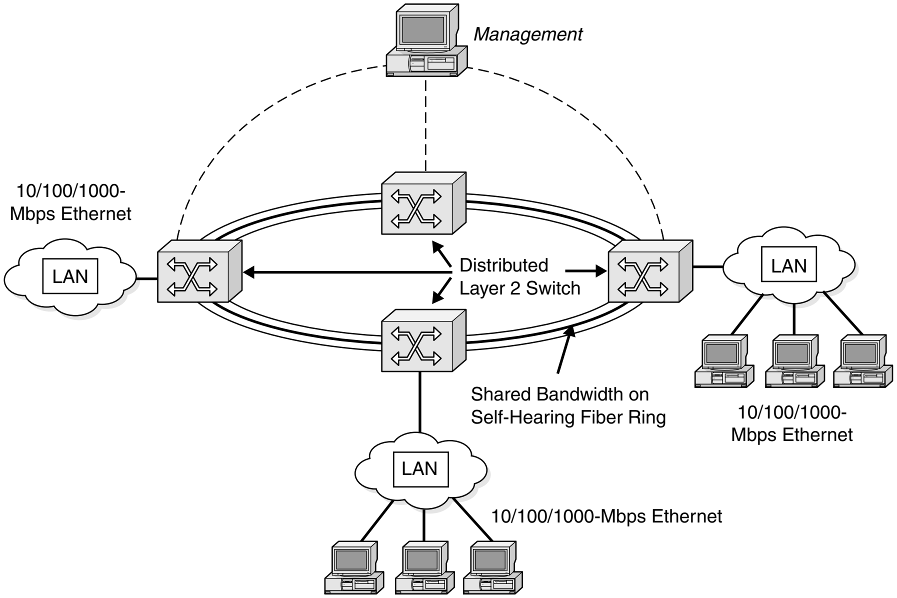


ring utilization, but it also enables new services such as usage-based billing. In the past, a multiservice operator had only two choices for metro delivery and transport: using an overlayed, routed IP network or mapping Ethernet into SONET frames. Because the first solution is built for data and therefore not optimized for multiservice, we must analyze the second SONET solution. With each 10-Mbps LAN service requiring an STS-1 bandwidth, an OC-48 (2.5 Gbps) metro ring has the capacity to carry 48 individual LAN circuits with approximately 1.9 Gbps of stranded bandwidth. This is not very efficient and it is one of the major reasons why incumbent service providers are not able to offer cost-effective LAN service over existing SONET networks. Virtual concatenation would enable the service provider to offer significantly more 10-Mbps LAN circuits. Given that a OC-48 metro ring has 1,344 VT1.5s, with each 10-Mbps circuit requiring 7 VT1.5s, the service providers can now deliver 192 LAN circuits in the same amount of bandwidth—an increase of more than 300 percent. It must be further understood that because each circuit is rarely operating anywhere near 10 Mbps, DBA can also be implemented. It is conceivable that an oversubscription of 4 could be engineered, meaning at any given time, 4 users could share an equivalent 10 Mbps of bandwidth. This results in an additional 576 10-Mbps LAN services, bringing the total to 768 users. This represents a network increase in efficiency of 1,500 percent.[1]

However, the picture is not as bleak (or bright, depending on your perspective) as just described. First, the stated inefficiency would only be manifest on the access ring, which typically can support 3 to 4 ADMs, and hence, only have 9 to 16 Ethernet circuits on the ring, assuming our realistic estimate of 3 TLS users per building. Secondly, it is well understood that some kind of Layer 2 grooming is needed at the CO/SN/POP. It is not desirable or feasible to solve the entire problem within the confines of just SONET (see Figure 4-15 for an example of Layer 2 grooming). In the past, ATM was used for Layer 2 grooming; now Layer 2 switches (such as a GbE switch) can be used to groom and repack the traffic. Furthermore, although 10-Mbps TLS was of interest in 1995,[15] even then these practitioners were already deploying 100-Mbps TLS. (The very idea embodied in the Fiber-Distributed Data Interface [FDDI] was to provide a MAN that had more bandwidth than the constituent subLANs.) If 100-Mbps Ethernet is mapped to an OC-3, a 2.5-Gbps ring would support 16 TLS connections, with a stranded bandwidth figure of 800 Mbps.

Wrapping up these considerations, the consensus appears to approach the bandwidth efficiency issue via a Layer 2 multiplexing technique, such as RPR as a better architecture and VLAN GbE as a lower-grade architecture, rather than using a Layer 1 VT-based approach.

Figure 4-15
Layer 2 grooming.



Next we describe another proposal. At the time of this writing, Nortel was advancing an alternative to a recently ratified standard for recovering Ethernet networks. The alternative claims to make better use of redundant switches, links, and bandwidth. The company has advanced its Split-Multilink Trunking Protocol (Split-MLT) technology as an alternative to the IEEE 802.1w Rapid Spanning Tree Protocol (RSTP) for reconverging metropolitan Ethernet networks. Split-MLT is targeted at carriers looking for more reliability of Ethernet services than subsecond recovery time. Split-MLT is based on the IEEE 802.1ab multilink trunking protocol, which is designed to enable backup links from a group of high-bandwidth trunks and traffic load balancing across those trunks. As such, Split-MLT enables service providers to configure redundant paths to a company or multitenant building without routing protocols such as Open Shortest Path First (OSPF) or the Routing Information Protocol (RIP). Split-MLT seeks to address three Ethernet reliability issues: link failure, node failure, and conversation preservation and maximum bandwidth utilization. For link and node failure, Split-MLT makes a pair of aggregation switches look like a single switch to the rest of the network. If one node fails, Split-MLT ensures that the remaining node has the destination MAC address information to forward traffic on behalf of the primary switch. This is done on a per-session basis, which helps ensure conversation preservation. If one of the links between the switches is lost, Split-MLT spreads sessions over the remaining working links; this makes more efficient use of available bandwidth by eliminating the wasted redundancy fiber that is used only if a failure occurs. Although Split-MLT might seem to be a more feature-rich alternative to RSTP, it still does not provide the 50-ms recovery of SONET, which is the benchmark that every service provider considering Ethernet is shooting for. The one-second recovery time of Split-MLT is the same for RSTP. RPR and next-generation SONET proposals such as Multiservice Provisioning Platforms (MSPP) are closer to what service providers are looking for in terms of Ethernet/SONET hybrids.[16]

In yet another approach, recent articles include the following: "There is a hot new product category emerging in the industry, but no one seems to know exactly what it is. Some call it next-generation SONET; others call it the Multiservice Provisioning Platform (MSPP). And since neither term has been clearly defined, vendors are making up their own definitions and then claiming to be market leaders."[17]

MSPPs combine transport, switching, and routing platforms into an integrated system with which service providers can offer bundled services' flexibly and ostensibly at a lower cost. By incorporating many service interfaces in one device, MSPPs may eliminate the need for extra devices to deliver intelligent optical services. Proponents also claim that MSPPs improve SONET's efficiency in transporting multiservice traffic. Next-generation SONET platforms, however, do much the same thing. MSPPs grew out of the next-generation SONET market. The MSPP category is really an umbrella category that includes next-generation SONET, and the lines are blurry between the two. Original next-generation SONET boxes were single-platform SONET ADMs, perhaps with an integrated crossconnect capability and Ethernet capabilities made possible through a service interface. Then MSPP startups started emerging and taking this model and adding other functions into it, such as Layer 2 switching and Layer 3 routing. Observers state that a device has to meet some criteria before qualifying as an MSPP. An MSPP device has to have a separate circuit for each service and a switching fabric for each type of traffic. A next-generation SONET device operates strictly at Layer 1: ADM, optical crossconnects, and a DWDM backend. MSPPs include switching and perhaps routing functions.

Some vendors, such as Astral Point, take the position that MSPP and next-generation SONET are two different categories. According to this vendor, MSPP is a shopping mall of services that can be used to access the carrier network; next-generation SONET only encompasses commonly occurring transport elements. Nortel also views next-generation SONET and MSPP as separate and distinct markets, and favors the former. According to Nortel, next-generation SONET works with the speeds that are out there and drives them into a single, cost-effective aggregation point. What we are seeing are these do-it-all-in-MSPP systems, but these systems are not granular, do not support higher speeds, and do not offer a smaller footprint. According to Cisco, the two terms define the same category. Cisco claims that its ONS 15454 metro optical system qualifies as both an MSPP and next-generation SONET platform. Appian Communications highlights a subtle difference between MSPPs and next-generation SONET devices—MSPPs tend to support packet services more efficiently than next-generation SONET platforms; next-generation SONET equipment has more port density, integrated ADM, grooming, and crossconnect capabilities predominantly for TDM. According to Coriolis, MSPPs typically integrate packet switching into their platforms and can handle ATM, Ethernet, and Frame Relay circuits. Ciena's next-generation SONET box, the MetroDirector, integrates grooming and switching elements, and provides basic SONET/TDM functionality in a higher-density rack space while also providing support for GbE and ATM.[17] Therefore, it seems that only the passing of time will truly define the differences between the flavors of the two evolving technologies.

End Notes

1. G. Lee, "SONET: A Viable Solution to the Metro Access Network Bottleneck?" Opticalbiznet.com, June 2001.
2. For example, note this quote from J. McQuillan, *Broadband Networking at the Crossroad*, NGN 2001 Promotional Materials, October 11, 2001: "Another important realization in 2001 is that the anticipated opportunities for new data-centric service providers have not materialized, leading to a much slower pace of technology adoption. Probably the four biggest reasons for a slowdown in network innovation are BellSouth, SBC, Qwest, and Verizon. And the fifth reason is Telcordia. With the failure of so many DLECs and CLECs, we have returned to a situation almost like the 1980s, when telecom was dominated by regulated, monopolistic carriers who derived the bulk of the revenues and profits from legacy services. Compounding the problem, Telcordia has a monopoly position as exemplified in the Operations Systems Modification of Intelligent Network Elements (OSMINE) process. ILECs will only buy from telecom equipment suppliers with element management systems that conform to

OSMINE. This series of tests can cost a vendor over $10M, and many startups simply cannot afford it. The result? Traditional vendors have regained the edge in building the public infrastructure. Breakthroughs in technology are taking a back seat to continuity of vendor relationships and operational procedures."

3. R. A. Barry, "Intelligent Optical Networking Year 2001: A Practical Perspective," NGN Proceedings, 2001.
4. The latter case is currently less popular than the former.
5. Nortel Networks, Vertical Systems Group, Dain Rauscher Wessel (DRW) estimates.
6. Jon Cordova, "A Trend That Won't End," in "The Analyst's Corner," *Primedia Business Magazine and Media,* September 28, 2001, http://industryclick.com/newsarticle.asp?Newsarticleid=236185&SiteID=3.
7. For data applications, this may actually be too long.
8. Venture capitalists need to be extremely careful with carriers offering $1,000 a month for broadband services. In our opinion, these prices are completely unsustainable and can only result in one of two outcomes within the span of two to three years: a merger/acquisition by another carrier that eventually will raise prices or a course exemplified by the DLECs/CLECs/RLECs/web hosting companies—an eventual meltdown.
9. Nortel White Paper, "An Optical Ethernet Business Case for Incumbent Local Exchange Carriers—A Nortel Networks Business Case," Publication 56049.25-09-01, September 2001.
10. The following quote (Nortel White Paper, "A Layered Architecture for Metro Optical Networks," 56018.25/08-00, August 2000) supports the following thesis: "While the economic values of Ethernet seem compelling, the issue is how QoS and SLAs can be enforced by the carrier. This has been one of the key propositions behind ATM versus Ethernet. The need to enforce QoS within SLAs was based on an assumption of a scarcity of bandwidth. While this assumption has been true in the past, today's bandwidth has been doubling every nine months, twice the rate of Moore's law for chip capacity. With prices of 2.5G and 10G products dropping rapidly coupled with metro DWDM products, the cost per megabit is dropping rapidly for metro optical services. Hence, the significant operational costs and complexity of enforcing QoS to customers can be resolved simply by recognizing that bandwidth is no longer a scarce resource. Rather than expending significant resources to micro-manage bandwidth, it is suggested to

advocate throwing bandwidth at the problem. Rather than QoS, it is recommended to enforce Class of Service (CoS) for most applications. The engineering of the network with significant bandwidth capacity will reduce the complexity of enforcing stringent QoS guidelines. Furthermore, any cost of provisioning additional bandwidth will be more than offset by a savings in operational costs. In today's tight market for network management resources, the savings in operational costs become the overriding critical path item versus incremental capital spending for bandwidth."

11. Although according to *The ATM & IP Report* (July/August 2001), the worldwide ATM switch market was $4.7 billion in 2001, this is but a mere drop compared with the estimated $150 billion spent annually on telecommunications equipment worldwide (just considering the sales numbers for Cisco, Lucent, Nortel, Alcatel, Tellabs, Marconi, Siemens, and so on).

12. Nortel White Paper, "Product/Service Information on OPTera Packet Edge System for Native Ethernet and IP Services over SONET Networks" Issue 1, 56132.16, March 2001.

13. D. Minoli, P. Johnson, and E. Minoli, *Ethernet-Based Metro Area Networks* (New York: McGraw-Hill, 2002).

14. Nortel White Paper, "Comparing Optical Ethernet Resilient Packet Ring to Packet over SONET," 56029.25/01-01, January 2001.

15. D. Minoli and O. Eldib, *Telecommuting* (Boston, Mass.: Artech House, 1995).

16. Terri Gimpelson, "Nortel's Alternative for Metro Ethernet Recovery," *Network World* (October 1, 2001).

17. Terri Gimpelson, "MSPPs, Next-Gen SONET Mixing Optical Signals," *The Edge,* October 9, 2001.

CHAPTER 5

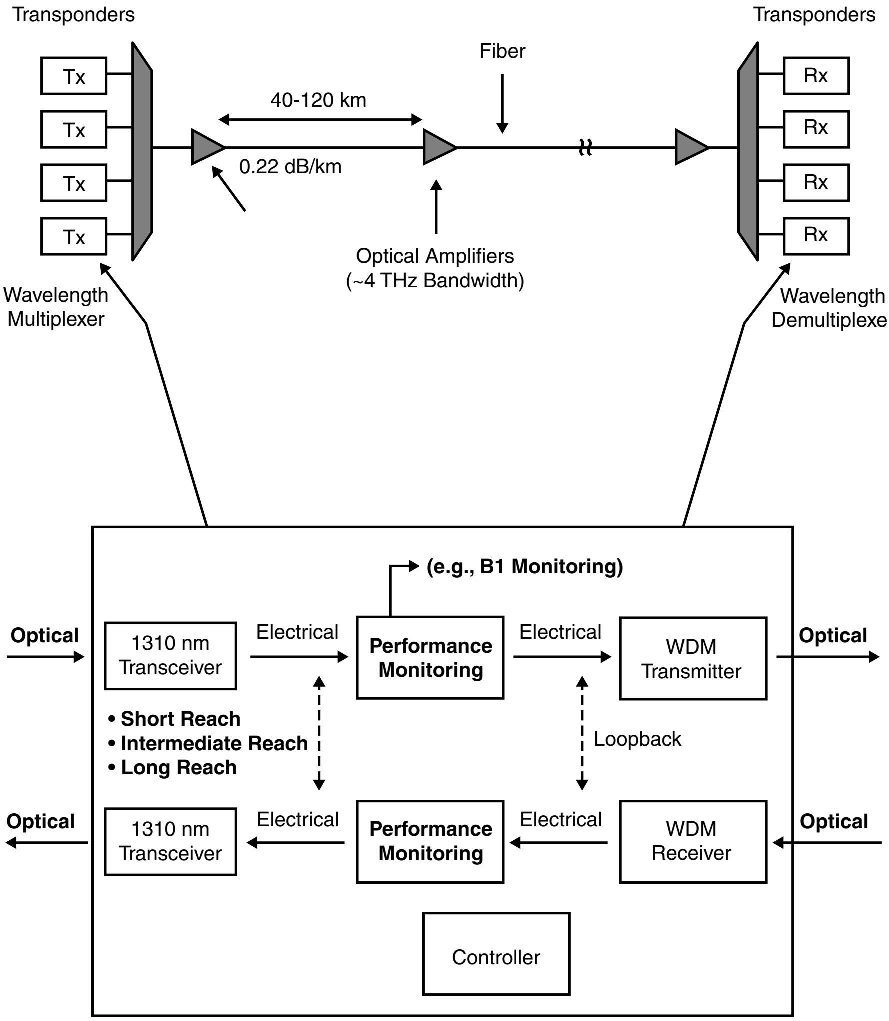
Dense Wavelength Division Multiplexing (DWDM) Systems

Wavelength division multiplexing (WDM) and high-capacity WDM systems called dense WDM (DWDM) have seen major deployment in the United States since the late 1990s. By using frequency-division multiplexing (FDM) techniques on a fiber transmission system, you can achieve fiber-optic pair gain where multiple signals (called *lambdas* [λ]) operating at different frequencies can each carry OC-48 (2.5-Gbps) or OC-192 (10-Gbps) channels between endpoints.[1] WDM is viewed as operating at the optical layer; Synchronous Optical Network (SONET) is viewed as an overlay onto this newly established layer. WDM multiplies the traffic-carrying capacity of existing fiber spans by combining two or more optical signals of different wavelengths on a single fiber pair.[2] DWDM coupler devices perform the actual multiplexing/demultiplexing of the different optical wavelengths. The multiplexing/demultiplexing of 2 to 1,000 channels has been demonstrated, with typical commercial systems in the field carrying 16 or 32 channels operating at OC-48 (2.5-Gbps) or OC-192 (10-Gbps) data rates. The International Telecommunications Union (ITU) has defined a set of standard wavelengths for manufacturers to use in DWDM in order to facilitate interworking between equipment. So far, this technology has been optimized for and has seen major deployment in long-haul networks based on very traditional cost-of-transmission-versus-cost-of-multiplexing economic trade-off considerations.[3]

Proponents of DWDM make the claim that acceptance of the technology will drive the expansion of the optical layer throughout the telecommunications network and enable service operators to exploit the bandwidth capacity that is inherent in optical fiber but that has gone largely untapped until now. At the same time, the widespread introduction of this technology could lead to a long-haul bandwidth glut and price disruptions,[4] and set expectations for price points at the metro access/metro core that may or may not be achievable. Vendors of DWDM systems and technology include Nortel, Cisco, Alcatel, Ciena, Lucent, JDS Uniphase, E-Tek, Photonic Integrated Research Inc., Kymata, Lightwave Microsystems, Bookham, Silkroad, and Nanovation to list a few. Many U.S. carriers have settled on DWDM at STM-16/OC-48 rates as their technology of choice for gaining more capacity at the long-haul level, but OC-192 (and soon OC-768 [40-Gbps]) systems are also in play. Sixteen-channel DWDM has been deployed throughout the carrier infrastructure, and 32-channel systems and above are also entering the field.

This chapter examines DWDM technology and its applicability, or lack thereof, to the metro access/metro core segment. We will focus on the economics of this candidate solution. Some of the underlying technologies for WDM/DWDM were discussed in Chapter 2, "Optics Primer." Here, we make only a brief reference to phased-array technology and claim that it is perhaps the most applicable technology to the metro access environment. The basic elements of a WDM system, including the transponders, are diagrammed in Figure 5-1.

Figure 5-1
WDM system and transponders



5.1 Opportunities and Application Scope

Because of its costs, DWDM is more suited to longer-reach applications at this time. This situation could change in the future if developers begin to grasp what the real requirements are in the metro access/metro core space. For long-haul applications, DWDM offers a cost-effective complement to OC-48/OC-192 time-division multiplexing (TDM) technology that is embodied in SONET/Synchronous Digital Hierarchy (SDH) systems. Assuming a decent-quality fiber plant, WDM methods enable the planner to maximize network capacity and address service scalability needs. In addition, because it defers—or eliminates—the capital outlays and long lead times associated with deploying new fiber cable, DWDM is a useful solution for high-growth routes that have an immediate need for increased bandwidth. Carriers that are building or expanding their long-haul networks could find DWDM to be an economical way to incrementally increase capacity, rapidly provision needed expansion, and "future proof" their infrastructure against unforeseen bandwidth demand. Network wholesalers can take advantage of DWDM to lease capacity rather than entire fibers either to existing operators or new market entrants. However, in aggregate, DWDM technology is not cheap. Systems can cost from $50,000 to $100,000 per lambda, particularly when redundant and long-reach applications are considered. Furthermore, these figures do not include the cost of the repeaters. Systems costing from $0.5 million to $1 million or more are not unusual. Power consumption and floor space requirements are usually high, and unless tunable laser systems are available, complex design/engineering and sparing disciplines are involved, requiring plug-ins that are specific to the given lambdas. Furthermore, future proofing may in fact imply that after a route upgrade, there is no need for additional equipment for several years; this is, therefore, bad news for equipment suppliers.

The proponents typically present diagrams similar to the right-hand side of Figure 5-2. At a glance, this seems to imply that DWDM trumps SONET because it requires only one fiber pair (quadrant D) instead of multiple fiber pairs (quadrant B). However, the true economics vary with demand and deployment situations. Assume that the carrier in quadrant A has a demand for increasing bandwidth. Clearly, its preferred migration path is to deploy DWDM, rather than the architecture of quadrant B, which would require additional fiber. On the other hand, assume that the carrier is in the situation depicted in quadrant B and already has several parallel systems deployed along with a demand for increasing bandwidth. A migration path

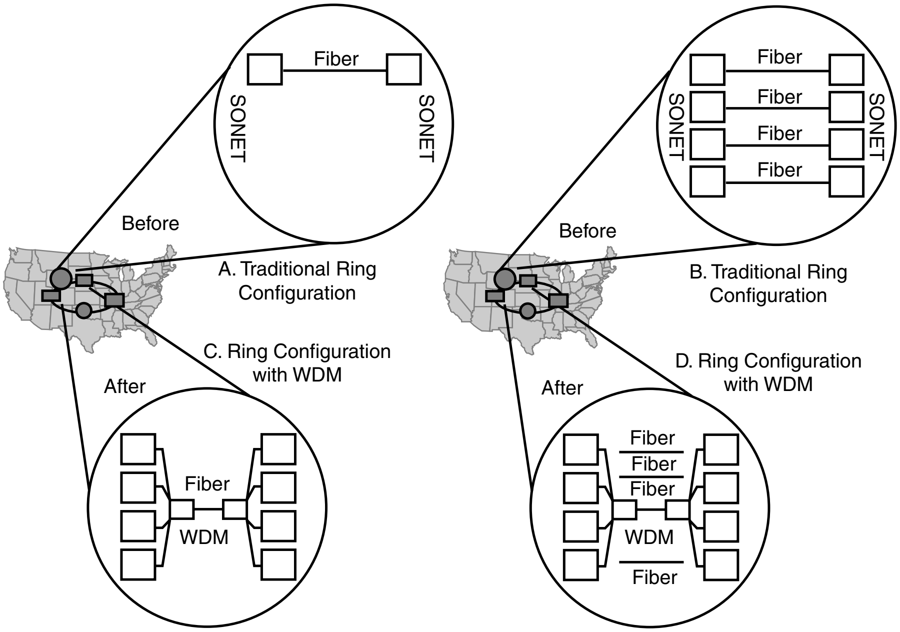


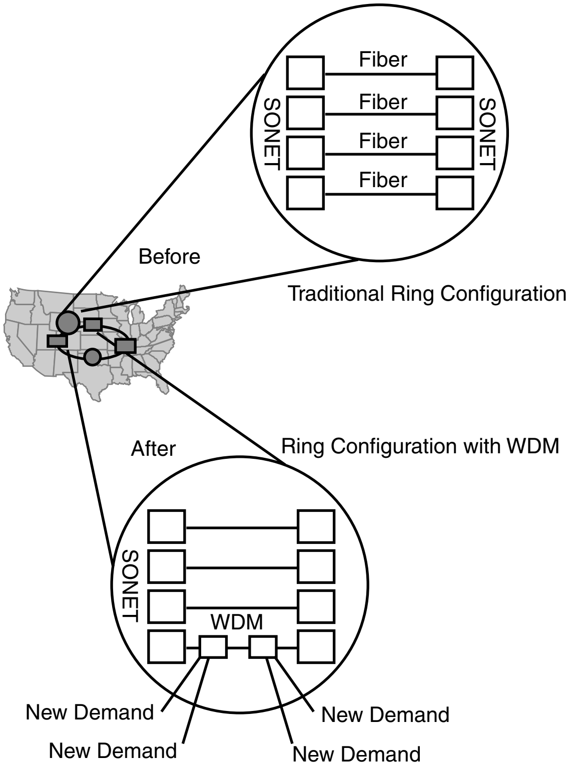
Figure 5-2
Design comparison: Left: increasing demand but with no multiple systems already deployed. Right: increasing demand but with multiple systems already deployed.

to DWDM, as specifically shown in quadrant D (the favorite final stage of proponents of DWDM), is not that optimal because it would leave a number of assets already owned by the carrier stranded. Hence, in this case, the migration path would probably be more like the one shown in Figure 5-3, and the dramatic saving implied by Figure 5-2 may or may not be forthcoming in this case.

DWDM is well suited for long-distance telecommunications operators that use either point-to-point or ring topologies. The availability of 16, 32, or more new transmission channels where there used to be 1 improves an operator's ability to expand capacity and simultaneously set aside backup bandwidth without installing new fiber. Proponents make the case that this large amount of capacity is critical to the development of self-healing rings. By deploying DWDM terminals, an operator can construct a protected 40-Gbps ring with 16 separate communication signals using only 2 fibers. However, unless there is a major underlying engine continuously driving the demand through the roof, this kind of technology is a *one-time upgrade*, which has obvious market sizing implications.

The pertinent question for the current discussion is whether DWDM has practical application in the metro space. There has been a lot of hype in the

Figure 5-3
Practical migration: increasing demand but with multiple systems already deployed.



recent past about metro DWDM. It is true that long-haul analog carrier techniques developed in the 1940s were later adopted by analog loop carrier techniques in the 1950s. The N1 and N2 carrier terminals supported 12 channels, whereas the N3 and N4 carrier terminals supported 24 channels. The long-haul transmission systems, however, supported 600 (L1), 1,860 (L3), 3,600 (L4), 10,800 (L5), and 13,200 (L5E) channels. Similarly, T1 (24 channels), T2 (96 channels), and T3 (672 channels at 44.736 Mbps) carrier systems developed in the early 1960s were later used in the 1970s for access in metropolitan digital loop carrier environments.[5] However, T4-NA (2,016 channels at 139.264 Mbps) and T4[6] (4,032 channels at 274.176 Mbps) were not used in loop applications because they were overkill.

The reality is that in its general form DWDM has some applicability to metropolitan environments, probably for a handful of point-of-presence-to-point-of-presence (POP-to-POP) rings.[7] If DWDM systems now in the market were redesigned to be optimized according to the requirements for metropolitan environments, there could be increased applicability. If the systems were redesigned to meet specific price points (see Section 5.2 “Overview of Technology”), their applicability would be enhanced. (Section 5.4, “Example of Optimized Edge Equipment” briefly discusses what some metro access equipment should look like.)

When the long-haul industry saw major retrenchments at the dawn of 2000, a number of optical vendors took the easy course of relabeling the equipment that they had developed for long-haul applications and pasted a Metro DWDM label onto the equipment while generating new marketing collateral. This was done instead of redeveloping equipment that is optimized and right sized for metro access/metro core applications from both a density and cost standpoint (as would have been more appropriate). It appears that, at least for the next few years, the opportunity for metro DWDM is somewhat limited. As noted, this technology may see some penetration in a metro core application, such as POP-to-POP rings, but it will probably see extremely limited application in the metro access segment (see Section 5.2). Analysis shows that these vendors have products that would cost from $190,000 per building all the way to $720,000 per building. This is totally untenable for metro access because carriers generally need to keep their in-building costs at no more than $20,000 or thereabouts. A price of $10,000 per dropped lambda ($20,000 for a redundant system with two lambdas) is desirable for metro access, particularly when the component cost of metro optics to build 8 or 16 channels of coarse WDM (CWDM), wide WDM (WWDM), or even DWDM is actually about $2,000 or less per lambda. In particular, if systems are designed in such a manner that each building needs to incur the full cost even though only one or two lambdas are dropped off at that building, then the equipment will see very limited penetration at the metro access level.

On the issue of component cost, phased-array wavelength division multiplexing devices (also called arrayed waveguide grating [AWG])[8] are emerging as important passive devices and may be the right kind of technology for metro access and possibly for some metro core applications in smaller-diameter tier-2 or tier-3 cities. New technology is available that can produce a 1×8 WDM chip (8 channels) for around $450, a 1×16 WDM chip for around $700, and a 1×32 WDM chip for $1,200. Component vendors may benefit from a sol-gen-based method that reduces manufacturing costs better than one order of magnitude when compared to other methods.

Sol-gen is a mature glass technology already used in several fields. Photonic hybrid active silica integrated circuit (PHASIC) is a Canadian-developed method that combines features of organic polymers and glass to produce a sophisticated hybrid material for integrated optics in a cost-effective manner. Companies can use advanced photomicrolithography to print optical circuits and devices directly into the hybrid glass. Literature is available[9] that documents the connection between molecular (chemical) hybridization, supramolecular hybridization, and functional hybridization (device- and system-level constructs). This literature describes how low-temperature hybrid sol-gel glasses (HSGGs) combine attractive features of organic polymers and inorganic glasses, giving rise to a new process for the fabrication of WDMs, couplers, power splitters, waveguides, and gratings. The process combines chemical synthesis and sol-gel processing with straightforward photomask techniques. HSGGs are material systems whose properties can be predictably tuned from properties that are closely related to organic polymers all the way to properties that are pretty much like conventional inorganic glasses. This creates a development opportunity for the chemist and material scientist, through appropriate design, to realize compositions that can be adapted to a range of optical devices. In academic lingo, some might say, "In combining broad spectrum molecular chemistry with physical processing, HSGG recommends itself as the material locus for resolving antagonisms that arise out of conflicting technical objectives in photonics."

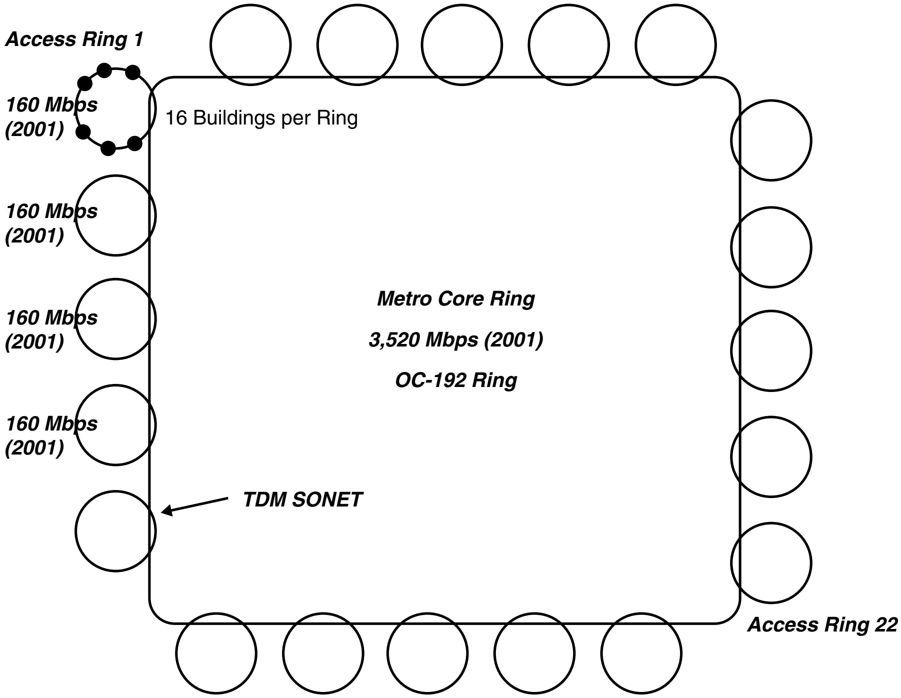
Unlike conventional/monolithic approaches, which make use of semiconductor materials, hybrid architectures for glass optoelectronics enable the development of a range of passive glass components such as couplers, splitters, waveguides, WDMs, mirrors, gratings, and lenses.[10] The five existing conventional glass fabrication methods all suffer drawbacks when compared with sol-gel techniques. These conventional methods are sputtering, thermal oxidation and nitridation, chemical vapor deposition (CVD), plasma-enhanced low-temperature chemical vapor deposition (PECVD), and flame hydrolysis deposition (FHD). Some vendors (such as PIRI, Siemens, and Kymata) use the FHD method. With this method, a hydrogen-oxygen nozzle flame burns gasses; the combustion of these gases produces a glass soot that is melted at ≈1,000°C. This process takes more than six hours and is repeated three times. The second coating step is followed by a series of coating and vacuum-etching steps. The PHASIC method, on the other hand, uses an HSGG; the components are made by spin- or dip-coating fluids at 100 to 200°C. Next, the devices are created by photolithography directly in the HSGG, avoiding vacuum film deposition and repetitive steps. One issue with passive devices, however, is power. After a few nodes, the original power is reduced appreciably; however, this issue is manageable in a metro access environment with relatively small (in diameter) rings. The phased-array technology is similar in concept to the passive optical network (PON) systems that are being discussed by the industry for residential use.

The best market research available indicates that today there are 21,000 office buildings in the United States on fiber, and the number is expected to increase to about 32,000 office buildings by 2005.[11] As an average, considering 60 major cities in the United States, there were 350 office buildings on fiber for each city in 2001 and will be 530 office buildings on fiber for each city in 2005. Further, assuming access rings with 16 buildings in each ring, this implies that as an average, there were 22 metro access rings for each city in 2001 and will be 33 metro access rings for each city in 2005. This would imply that a single metro core DWDM system (with 32 lambdas) interconnecting these access rings at the various citywide POPs would suffice in 2001, and 2 metro core DWDM systems (each with 32 lambdas) interconnecting these access rings at the various citywide POPs would suffice in 2005 (we are not assuming for this discussion that a DWDM is needed in the access ring). Even assuming that there were 5 carriers in each city sharing (or doubly connecting) these buildings, this would equate to 5 metro core DWDM systems for each city in 2001 (that is, 300 systems nationwide) and 10 metro core DWDM systems in 2005 (that is, 600 systems nationwide). Assuming a cost of $2 million per citywide system to connect the various POPs in question (for example, $250,000 per node with 8 POPs in a city), this would equate to an embedded market of $600 million in 2001 and $1.2 billion in 2005 (or about a $150 million increment per year). This estimate is much lower than the estimate offered by many proponents, except for the number provided by CIR.[11]

Keeping in mind the average of 22 metro access rings per city in 2001 and 33 metro access rings per city in 2005, we claimed in Chapter 1, “Advances and Opportunities in Next-Generation SONET and Other Optical Architectures,” that the average per-building bandwidth was 10 Mbps in 2001 and would be 73 Mbps in 2005. Figure 5-4 depicts a generalized 2001 design that has 22 access rings (these being well served by an OC-48 SONET arrangement) supported by 1 OC-192 SONET metro core ring. The 2005 situation needs more access rings as well as more bandwidth on each of the access rings. The 33 access rings can be served by OC-48 SONET systems. The metro core ring can either be SONET OC-768 or a WDM ring supporting 4 OC-192 lambdas (see Figure 5-5). From a traffic-engineering standpoint, this discussion shows that perhaps one WDM system per city (per carrier) would suffice in a typical city. (Of course, some cities are larger, and other cities are smaller.)

Reexamining the market scope of long-haul-based DWDM, you should note that because the traditional voice market has not grown more than a few percentage points a year for the past couple of decades, the growth in long-haul demand must come from either Internet applications or from

Figure 5-4 Typical city arrangement in 2001.



wireless networks trunking. Table 5-1 captures the bandwidth requirement assuming an 80 percent per year growth, which was the best estimate of the growth of the Internet at the end of 2001. The interesting thing to note is that if a carrier had an OC-48 between two points and the carrier upgraded that link with a DWDM system with 16 channels, then there would be no need to deploy other hardware for 6 years. If the carrier upgraded the system with a 32-channel solution, then there would be no additional need to deploy other hardware for 7 years. Clearly, this is not good news for long-haul DWDM manufacturers.

Taking this analysis a step further, assume that there are 20 tier-1 network service providers (NSPs) in the United States. (The actual number is probably smaller.) The 80 percent aggregate growth of the Internet would then translate into a 4 percent growth per individual NSP, assuming an even distribution. Because of the inefficiencies associated with granularity, we actually assume 2.5 times the growth with each NSP; namely, we assume a 10 percent growth per year per NSP. Further, we assume that a typical national NSP network approximates the one shown in Figure 5-6, which has 9 nodes and 12 links.

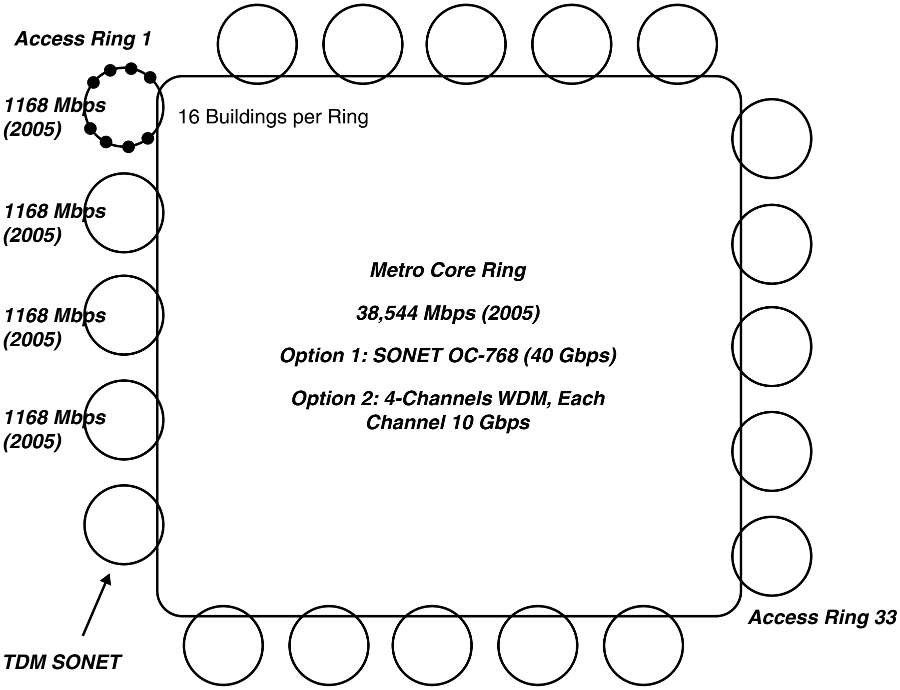


Figure 5-5 Typical city arrangement in 2005.

Table 5-1 Upgrade Now and Do Not Touch for Seven Years (80 Percent Yearly Growth in Demand)

Year	Demand (in Bandwidth Units)
Year 1	1.00
Year 2	1.80
Year 3	3.24
Year 4	5.83
Year 5	10.50
Year 6	18.90
Year 7	34.01

Assuming that all 20 NSPs purchased 24 DWDM systems (two per link) on day 1, it would equate to a market of \$1.6 million×20×24=\$768 million (assuming 16 lambda systems and \$100,000 per lambda). As Table 5-2 shows, once an NSP makes a one-time networkwide upgrade, the need for further upgrades does not appear until the far future.

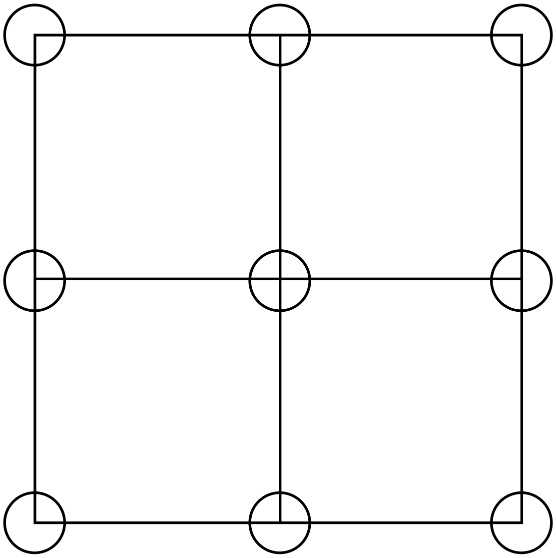


Figure 5-6 Synthetic NSP national network.

Table 5-2 Upgrade Now and Do Not Touch for Many Years (Allocated 10 Percent Yearly Growth in Demand)

Year	Demand (in Bandwidth Units)
Year 1	1.00
Year 2	1.10
Year 3	1.21
Year 4	1.33
Year 5	1.46
Year 6	1.61
Year 7	1.77
Year 8	1.95
Year 9	2.14
Year 10	2.36
Year 11	2.59
Year 12	2.85
Year 13	3.14
Year 14	3.45
Year 15	3.80

Table 5-2 cont.

Upgrade Now and Do Not Touch for Many Years (Allocated 10 Percent Yearly Growth in Demand)

Year	Demand (in Bandwidth Units)
Year 16	4.18
Year 17	4.59
Year 18	5.05
Year 19	5.56
Year 20	6.12
Year 21	6.73
Year 22	7.40
Year 23	8.14
Year 24	8.95
Year 25	9.85
Year 26	10.83
Year 27	11.92
Year 28	13.11
Year 29	14.42
Year 30	15.86
Year 31	17.45

Even in long-haul applications, design considerations aimed at optimizing the cost profile are not always straightforward. In particular, TDM-only solutions supporting increasing speed have kept pace with a number of advances in WDM technology during the mid to late 1990s, at least for medium-size trunking applications (up to 10 Gbps). For example, a TDM-based solution has only one 10-Gbps SONET terminal, as shown at the top of Figure 5-7. A WDM system that transports an aggregate capacity of 10 Gbps requires four 2.5-Gbps terminals in addition to a WDM terminal per end (shown at the bottom of the figure). Because TDM technology has typically quadrupled its capacity for a cost multiplier of 2.5, the 10-Gbps solution appears to be more cost effective. However, if the TDM system also requires four 2.5-Gbps terminals to provide the first stage of multiplexing, the 10-Gbps solution might actually be more costly. (Note that if the

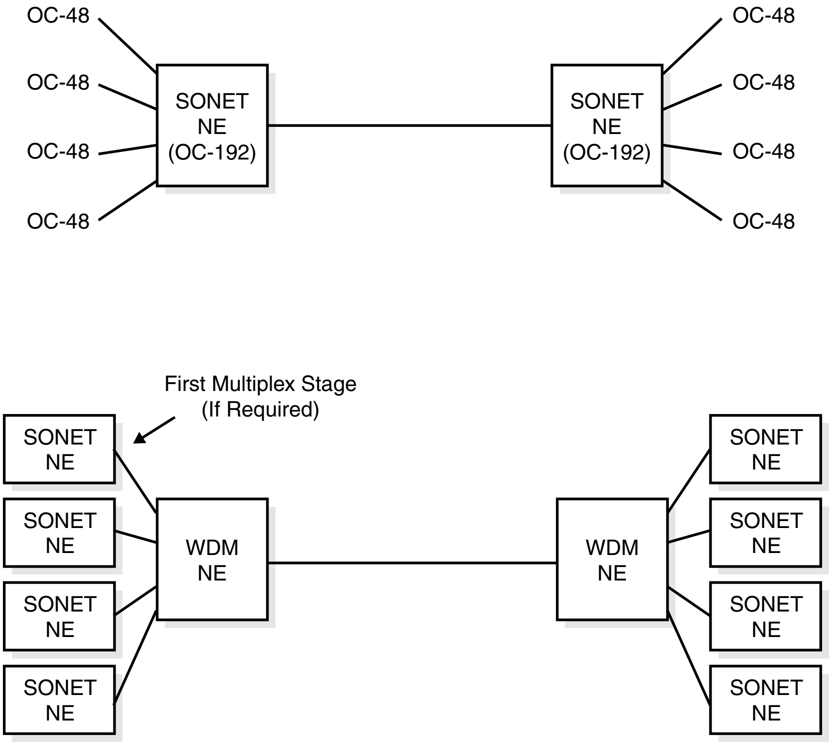


Figure 5-7 DWDM arrangement.

2.5-Gbps terminals are already in the network, they represent a sunk cost and might not be included in the cost analysis.)

Before optical line amplifiers (OLAs) were developed and deployed, higher-speed TDM-based systems were more cost effective than WDM because the TDM systems allowed multiple lower-speed electronic regenerators at one point in the network to be replaced with a single higher-speed regenerator at that point in the network. Originally, this was not the case with the WDM design (as implied in Figure 5-7). The introduction of OLAs with the capability to amplify the entire ITU grid frequencies (which reside in the erbium-doped fiber amplifier [EDFA] passband) simultaneously enables multiple lower-speed electronic regenerators at a site to be replaced with one optical amplifier, as shown in Figure 5-8, making WDM more cost effective in this context.

Having established a reality check on the baseline demand and market for this technology in both the long haul and metro space to address the industry hype, we can now proceed with a more technical discussion.

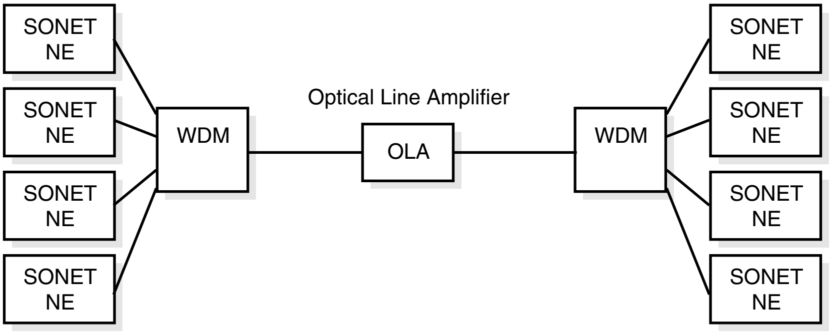


Figure 5-8
OLA.

5.2 Overview of Technology

5.2.1 TDM Evolution

Going back at least half of a century, network designers have looked for simple, straightforward methods for the migration of networks to higher performance levels. Throughout the 1950s and 1960s, first analog and then digital carrier systems of ever-increasing capacity were being developed. A carrier system is a multiplexed transmission system supporting multiple channels. Fiber-optic (lightwave) systems have been deployed since at least 1979 (one such system that operated at DS3 rates was deployed in Trumbull, Connecticut, in October 1979). As early as 1983, 90-Mbps systems (FT3C) were deployed. In the mid to late 1980s, systems supporting 400 to 565 Mbps emerged, followed by systems operating at 800 to 1,100 Mbps. In the early 1990s, systems operating at 1,600 to 2,200 Mbps (1.6 to 2.2 Gbps) emerged. Finally, SONET/SDH systems emerged supporting up to 2.5 Gbps (OC-48); 10-Gbps (OC-192) systems have been introduced in the past couple of years, and 40-Gbps systems are under development. These advances are based on TDM technologies, where the clock driving the on/off intensity modulation (IM) light process is driven at a higher pulse rate.

At the limit of how short the light pulse can be made, there are at least two other methods to increase carrying capacity on the fiber waveguide: coherent transmission, which is a variant of full FDM, and WDM, which is in effect a space-division multiplexing method where distinct light beams are modulated with IM methods, as shown in Figure 5-9. Although coherent methods offer perhaps the best long-term opportunity for maximum

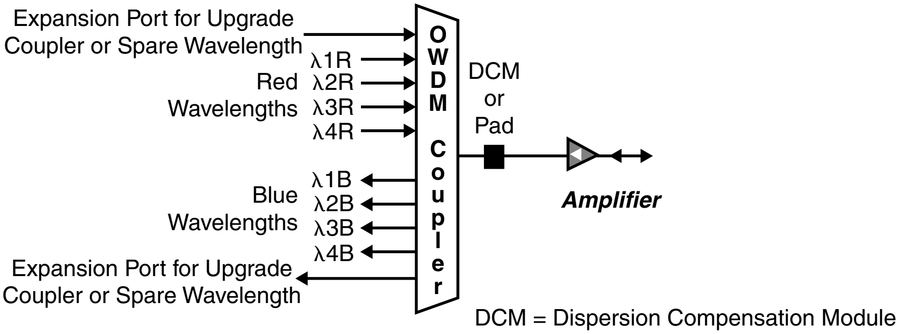


Figure 5-9 Basic DWDM multiplexing method.

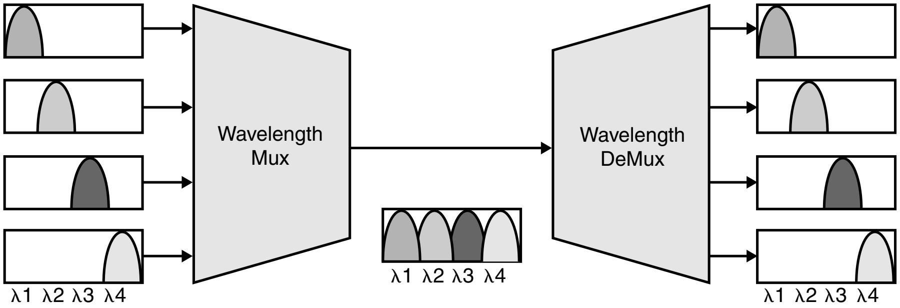


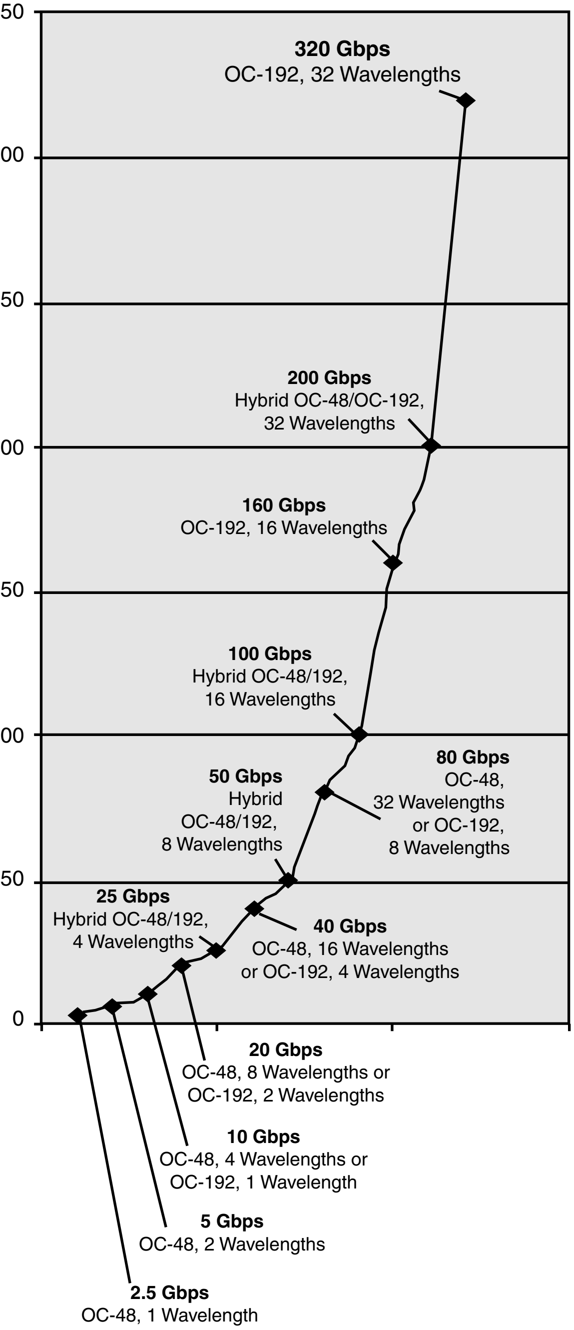
Figure 5-10 Pictorial view of WDM.

throughput, they currently suffer from technical, engineering, and cost drawbacks. Therefore, WDM methods have enjoyed significant market penetration in the long-haul segment in the United States and abroad. Figure 5-10 provides a pictorial view of WDM. Figure 5-11 depicts the progression of throughput in WDM systems over the past few years.

The following list depicts the evolution of transponder technology in the recent past:[12]

- **Capacity** OC-48 → OC-192 → OC-768 → OC-*n*
- **Spacing** 400 GHz → 200 GHz → 100 GHz → 50 GHz → 25 GHz
- **Bit rate transparent**
  - Issues—Jitter problem and performance monitoring
- **SONET rate transponder**
  - Jitter compliant
  - Issues—B1/J0 monitoring, fault isolation, and misconnections

Figure 5-11 Bandwidth improvement for long-haul systems over time.



- **Section-terminating equipment (STE) transponder**
  - Full-compliant SONET STE
  - Software configurable (B1/J0)
- **Line-terminating equipment (LTE) transponder**
  - Muxing and add/drop
  - OC-48 and OC-192
- **Path-terminating equipment (PTE) transponder**
  - Mapping new services (such as Gigabit Ethernet [GbE]) into SONET frames

The proliferation of bandwidth-intensive applications has steadily driven the demand for broadband connectivity and transport. DWDM increases the capacity of the embedded optical fiber base by adding pair-gain methods to the transmission channel. A DWDM optical backbone gives carriers the flexibility to expand capacity in any portion of their network, thus addressing specific problem areas that are congested due to high bandwidth demands. However, any network planner knows that as you move closer to the core, the need becomes more intensive and concentrated, whereas at the edges of the network, the need is less intensive and concentrated. Therefore, the long-haul core is a more natural habitat for DWDM technology than the metro access and metro core space.

In long-haul networks, the primary value provided by DWDM in combination with OLAs is the cost-effective transmission of high-aggregate bit rates over large distances on a single fiber pair. The large distances in long-haul networks make deploying new fiber challenging. Long-haul carriers have traditionally been the first to embrace advances in transmission technologies to gain additional capacity while leveraging their existing equipment.[13]

The versatile nature of DWDM technology enables Asynchronous Transfer Mode (ATM), Internet Protocol (IP), and SONET networks to share the same optical core (see Figure 5-12), but possibly with some bandwidth inefficiencies. (For example, yes, you can map a GbE link directly over a lambda, but you will be dedicating a channel to support 1 Gbps where you could otherwise derive 2.5 Gbps/OC-48 or 10 Gbps/OC-192.) DWDM provides a means for carriers to integrate diverse technologies of their existing networks onto one physical infrastructure. DWDM systems are bit-rate and format independent and can accept any combination of interface rates and formats, although SONET tributaries are the most common. DWDM provides a grow-as-you-go infrastructure that allows for a 10- to 100-fold capacity expansion as new services increase demand for bandwidth. The grow-as-you-go approach is implemented by adding the lambda plug-ins

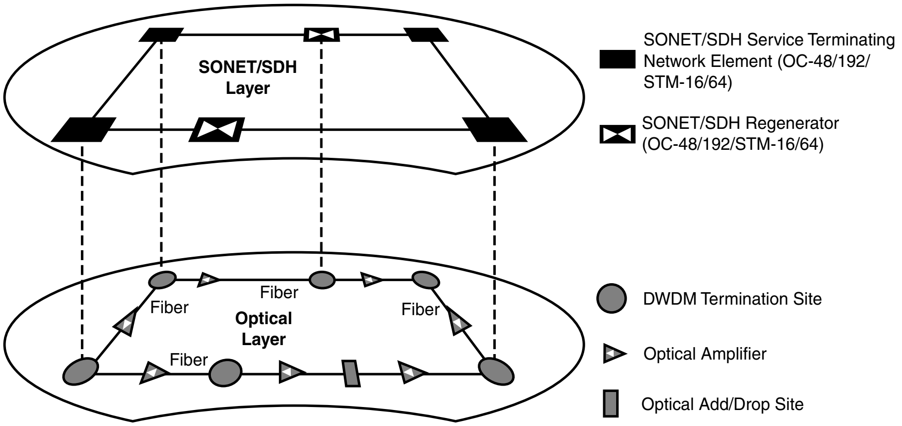


Figure 5-12 SONET versus optical plane.

over time.[14] High-end commercially available DWDMs can deliver up to 320 Gbps of capacity; these are 32-channel systems with each channel operating at 10 Gbps (OC-192). Optical technology under development at the commercial level can currently provide 10 Gbps on a single lambda and 160 lambdas over a single fiber pair for long-distance networking services (more lambdas are achievable in laboratory systems). However, many observers now believe that there was an overestimation of the demand in the late 1990s and early 2000s, at least in the long-haul and Internet segments of the market, and that those very large systems will not be needed in the near future.

5.2.2 EDFA Role in DWDM

In WDM, multiple wavelengths of light carry the data channels to an optical multiplexer where they are combined for transport over a single fiber pair, as shown in Figure 5-13. Essential to WDM technology is the optical amplifier, particularly for long-haul applications because they usually don't have much applicability in the metro space (see Figure 5-14). Optical amplifiers maintain continuity by amplifying the multiple signals simultaneously as they travel down the fiber. The initial amplifier technology has EDFA. Now new extended-band amplifiers are being studied and developed. These amplifiers carry the signal to optical demultiplexers where they are split into the original channel. Semiconductor optical amplifiers have been proposed for use in optical crossconnects (OXCs) for wavelength conversion in the optical domain and the amplification of a large number of ports. Support technologies include optical components (mux/demux

Figure 5-13 Basic DWDM transmission arrangement.

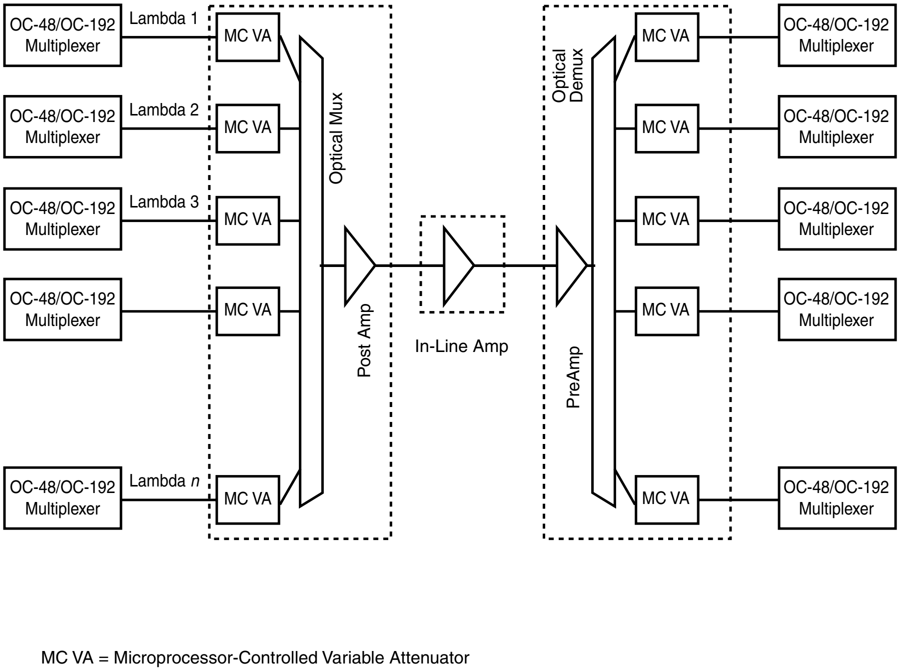
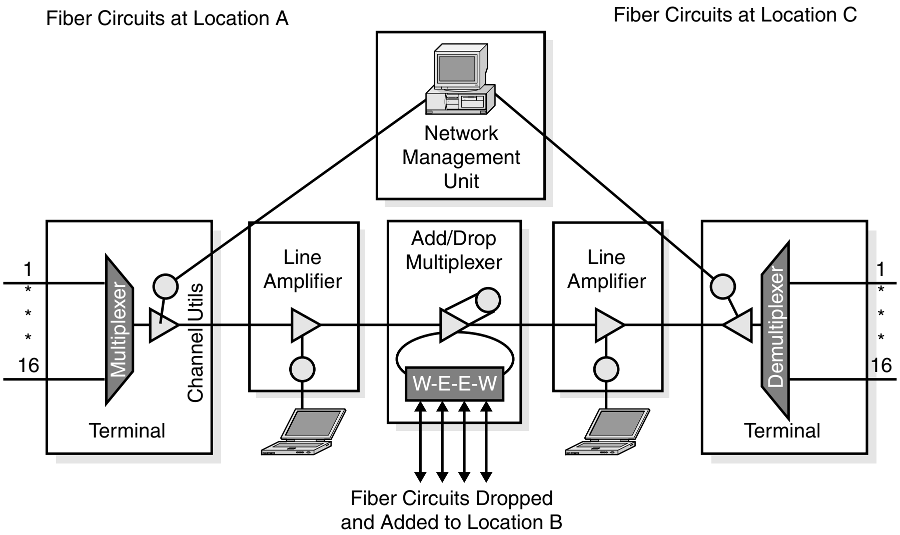


Figure 5-14 Basic DWDM block diagram.

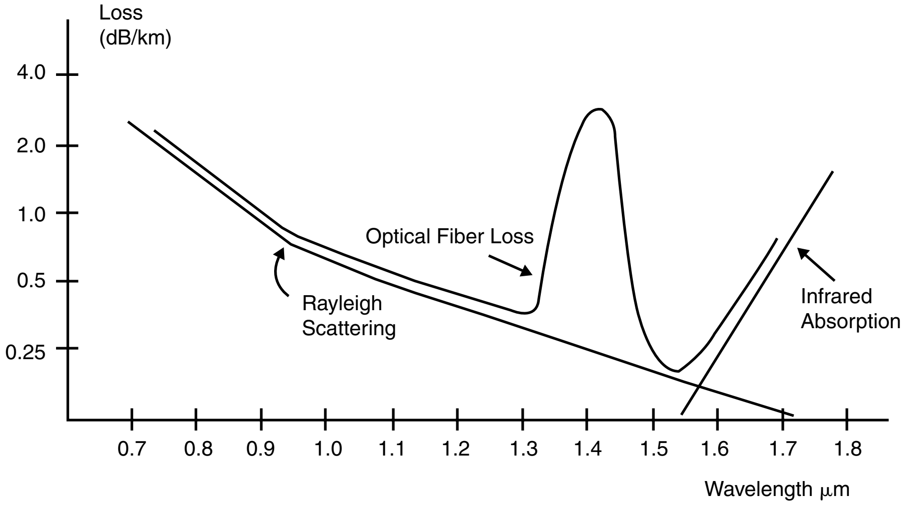


devices and dispersion compensation) and design enhancements to overcome the fiber nonlinearities that cause signal distortion in transmission. DWDM wavelengths fall within two bands: a blue band between 1,527.5 and 1,542.5 nm and a red band between 1,547.5 and 1,561 nm. Each band is dedicated to a particular direction of transmission. See Figure 5-15 for a view of the channel transmission characteristics of single-mode fiber cable.

Looking at Table 5-3, we can see the distances on different types of fiber that EDFA enables. EDFA also enables the following:

- Amplification in the optical domain (no optical-to-electrical-to-optical [O-E-O] conversion required)

Figure 5-15 Transmission characteristics of common fiber.



Rayleigh scattering: 0.15 to 2.3 dB/kb in 0.85 μm to 1.55 μm region OH ion absorption: 0.01 to 5 dB/km (silica impurity)

Infrared absorption: Sharply increases beyond 1.60 μm region Waveguide imperfections: up to 0.1 dB/km (other scattering problems: Brillouin scattering and Raman scattering)

Table 5-3

Supportable Distances Amplifier to Amplifier

Fiber Type	OC-48 Applications	OC-192 and Hybrid OC-48/192 Applications
NDSF	700 km	320 km
NZDSF	700 km	450 km
DSF	700 km	400 km

- Amplification of multiple signals simultaneously
- Format independence (analog, digital, pulse, and so on)
- A 20- to 30-dB gain and output power
- Bit-rate independence when designed for the highest bit rate
- Current bandwidth: 30 nm in the 1,530- to 1,560-nm range (80 nm is possible in the future)
- High bandwidth that can be extended with improved designs

WDM amplifier technology has improved significantly in the past few years. It is now mature and is used in many networks. These amplifiers support the transmission of multiple OC-48 (2.5 Gbps), OC-192 (10 Gbps), and evolving OC-768[15] signals over conventional and newer nonzero dispersion-shifted fiber (NZDSF). Typical repeaters enable 32 channels to be inserted and spaced 0.8 nm apart in field-deployable systems. You need flat gain performance in the bandwidth of the amplifier as well as high power. High power is critical in long-haul applications where you need to transmit over long distances. Figure 5-16 is an example of a large DWDM system (long haul). High-end DWDM systems are designed with extended reach optics support for the major fiber types in use today including

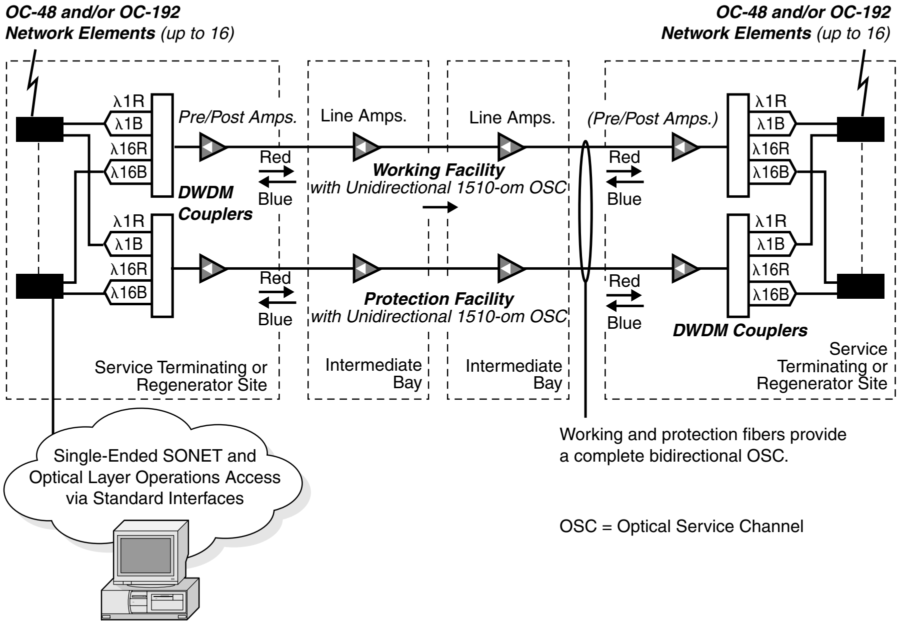


Figure 5-16 Example of large DWDM system (long haul).

nondispersion-shifted fiber (NDSF), dispersion-shifted fiber (DSF), and NZDSF (for example, Lucent TrueWave™ Classic/Plus and Corning SMF-LS™/LEAF™). (Fiber types were discussed in Chapter 2.) Long-haul systems typically need large cabinets, as depicted in Figure 5-17. This makes the systems expensive from an equipment, power, environment, space, maintenance, sparing, design, and troubleshooting standpoint, limiting their applications in the metro segment. Long-haul applications require periodic regeneration to ensure sufficient signal strength and quality at the receiving end. This results in substantial repeater bays, as shown in Figure 5-18, adding to the total transmission system cost. Although the cost of one

Figure 5-17 Typical DWDM equipment rack.

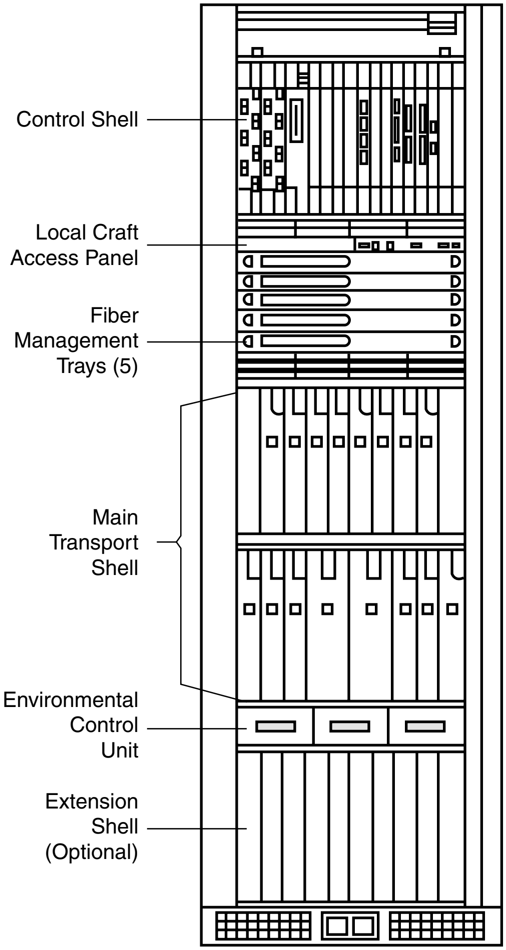
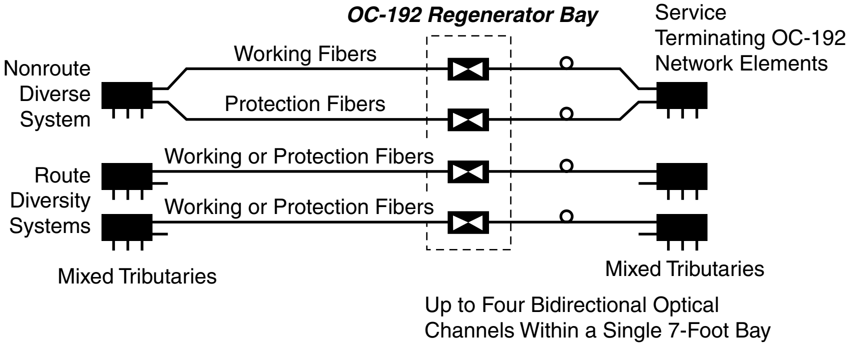


Figure 5-18
Regeneration bay.



Four-Channel Regenerator Configuration Serving Mixed Nonroute Diverse and Route Diversity Systems

regenerator may appear to be small, these costs can dominate when multiplied over the total length of a route, especially for heavy routes where there might be several systems contained within the same fiber sheath.

5.2.3 Technology Evolutions in DWDM

In Figure 5-19, we see a typical evolutionary environment for DWDM systems. The top portion shows a baseline transmission route that consists of four parallel SONET systems. As the need for expansion arises, WDM and TDM technology can be considered to reduce the number of fibers used in this application. WDM technology has traditionally been considered more expensive than TDM. One reason is that a WDM solution requires a separate SONET terminal for each channel in addition to the WDM terminal because of the embedded electronic regenerators, as shown at the bottom of Figure 5-19.[13] One way to address this issue would be to retire these electronic systems, assuming that they were fully depreciated, and replace them. However, carriers that are bottom-line positive in financial metrics are not fond of retiring working equipment.

Early deployments of WDM (in the early 1990s) were based on wideband technology. Figure 5-20 illustrates how wideband WDM (WWDM) doubles the capacity of fiber plant by optically coupling the outputs of two terminals in a fiber-optic transmission system (FOTS); one terminal operates in the 1,310-nm range and the other operates in the 1,550-nm range. Although this is a cost-effective solution for applications with restricted reach, wideband WDM systems, which tend to consist of a little more than an optical

Figure 5-19
Economic advantages of WDM.

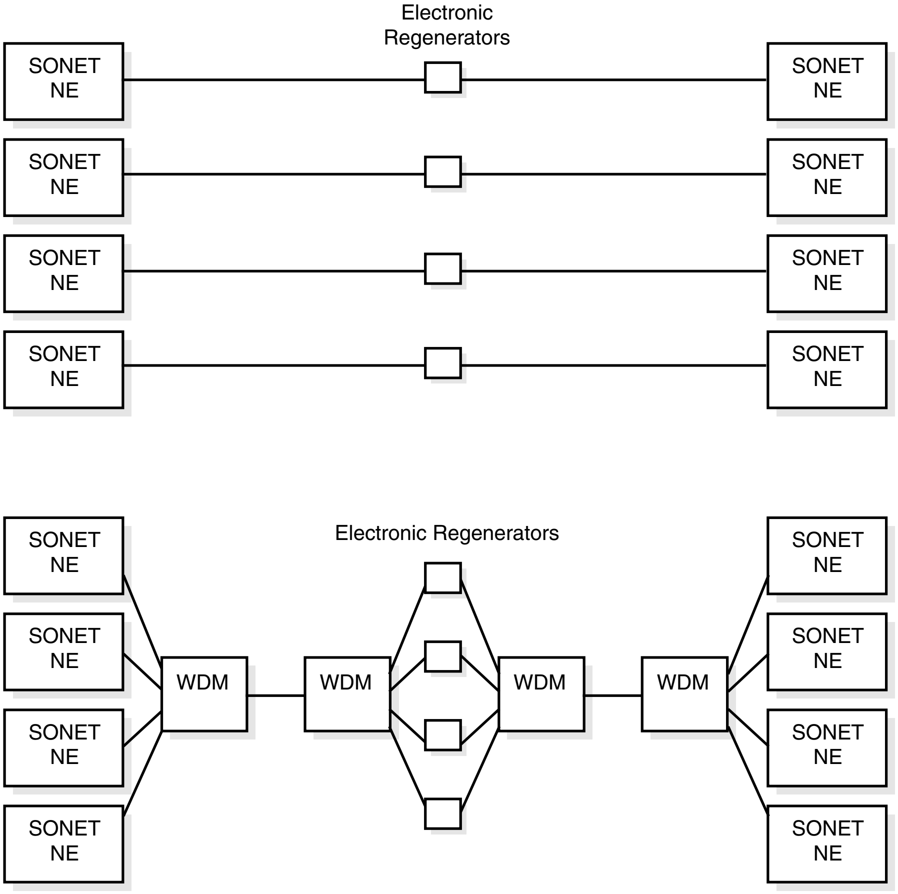
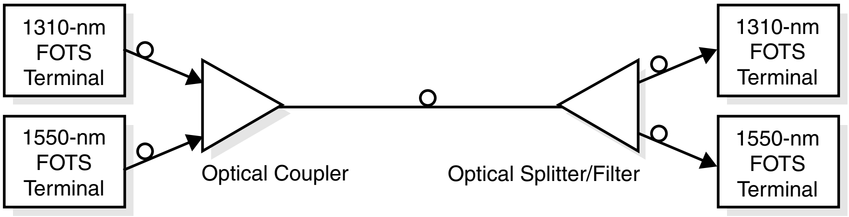


Figure 5-20
Original WDM system.



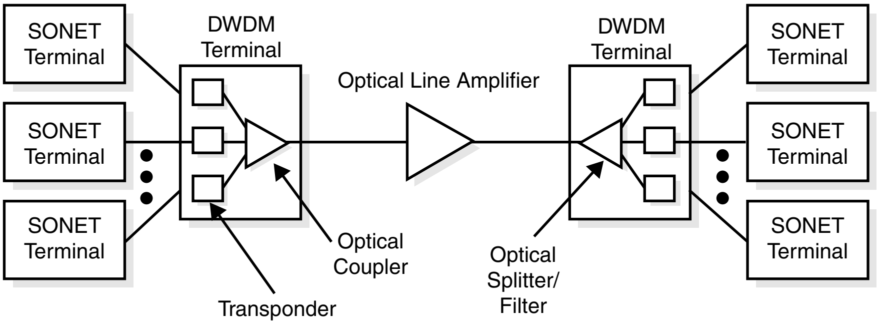
coupler and splitter, suffer from the absence of maintenance capabilities and scalability for long-haul applications.[13] However, if they are priced correctly, these WWDM or CWDM solutions, perhaps with some added features, may be applicable to metro applications.

More recently, narrowband or DWDM systems have been deployed. With DWDM, the outputs of 2, 4, 8, 16, or more SONET terminals are optically multiplexed into one fiber pair, as shown in Figure 5-21. As noted earlier, the wavelengths used are all within the 1,550-nm range to enable optical amplification using EDFA technology to maximize reach (stated differently, the EDFA operates in this optical region). The wavelength of the optical output from a SONET network element (NE) is roughly centered at 1,310 or 1,550 nm and has an approximate tolerance of ±20 nm. Such wavelengths would have to be so widely spaced that after optical multiplexing only a few would be within the EDFA passband. Therefore, DWDM systems must translate the output wavelength from the SONET NE to a specific, stable, and narrow-width wavelength in the 1,550-nm range that can be multiplexed with other similarly constrained wavelengths. The translation device is sometimes called a *transponder*.

The original DWDM systems were optimized for long-haul interexchange applications; therefore, these WDM systems support mainly point-to-point or linear configurations. Although these systems provide fiber capacity relief, the management of add, drop, and pass-through traffic must be done manually. Current-generation DWDM products can perform linear add/drop multiplexer (ADM) functions in long-haul or interexchange applications. These products also support ring topologies; ring topologies increase the reliability and the survivability of the user-level connection. Products that support ring and/or mesh configurations are more suited to metro applications.

However, current-generation systems still lack the full flexibility and affordability that is required in metro applications. Because the wavelengths that are added and dropped are fixed, there is no mechanism to increase the add/drop capacity without physically reconfiguring the system.

Figure 5-21
Simple WDM system.

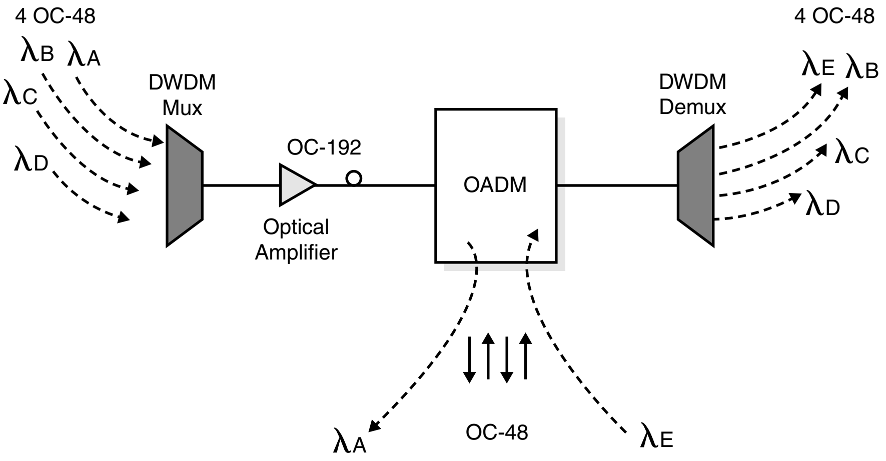


Networking in a TDM network is provided by multiplexers, such as traditional linear and ring ADMs and crossconnects. These NEs must convert the optical signal to an electrical signal before performing their specific function and then convert the result back to an optical signal for transmission on the next span. This O-E-O conversion is a major cost component of these network elements.

Photonic networking is the next-generation application of WDM technology. Future systems providing end-to-end photonic networking will have the capability to manage individual nonfixed wavelengths utilizing tunable lasers. Optical add/drop capability offers the ability to add or drop wavelengths from a single fiber without using a SONET terminal, as shown in Figure 5-22. Figure 5-23 shows the traditional approach, which entails using SONET terminals (transponders) at the intermediate drop-off point. Additional flexibility will come from crossconnect capability at the optical layer, combined with DWDM and optical add/drop multiplexer (OADM) functionality (see Figure 5-24). Photonic networking achieves multiplexing and crossconnecting strictly in the optical domain, significantly reducing the number of O-E-O conversions and the cost of the network. However, full photonic networking is not available today at the commercial level, and not all of the network elements discussed here were available or broadly deployed in early 2002.

Tracking the technology evolution, OC-192 SONET systems have emerged recently that can transport an OC-192 (10 Gbps) on one fiber pair. These can be deployed discretely (as a simple SONET/SDH system) or in

Figure 5-22
OADM.



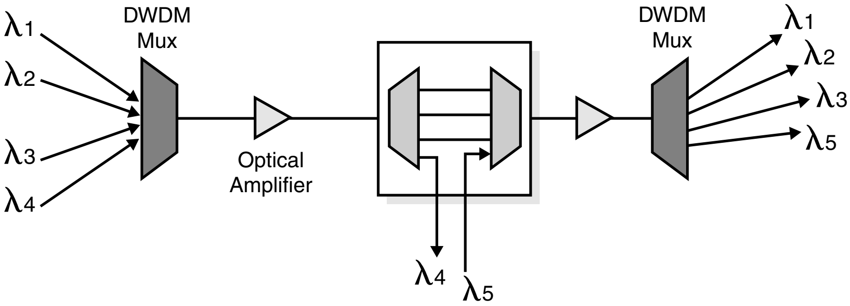


Figure 5-23 ADM with back-to-back transponders (O-E-O).

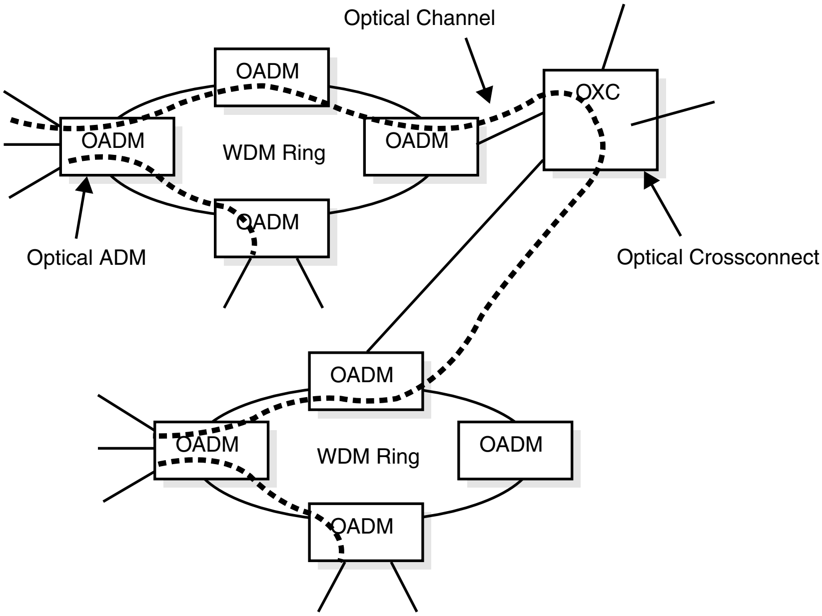


Figure 5-24 All-optical transmission system.

conjunction with a WDM system. There are a number of issues, though, that impact TDM-based data rates at 10 Gbps, including

- **Polarization mode dispersion (PMD)** Caused by a combination of intrinsic effects (birefringence and mode coupling) and extrinsic effects on cable (bends and twists). Studies show, though, that transmission up to 250 miles is possible at OC-192.
- **Stimulated Brillouin Scattering (SBS)** Limits power that can be injected into a single-mode fiber. As more OC-48/OC-192 systems are placed over a fiber, the output power of the EDFAs can cause SBS

problems. Techniques have recently been developed to boost the SBS threshold.

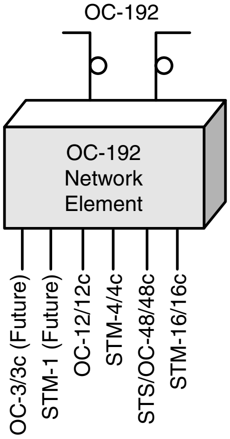
- **Self-phase modulation (SPM)** Introduces chirping and in turn interacts with fiber dispersion to cause pulse broadening or compression (depending on the dispersion profile of the fiber). This issue is also being studied with the goal of finding workable solutions.
- **Four-wave mixing** Limits multichannel transmission on DSF that has its zero dispersion in the EDFA bandwidth range. Monitoring input power levels into the EDFA can address this problem when conventional fiber is used for transmission. New fiber designs, including NZDSF, have been introduced to address the issue.

The majority of embedded fiber in long-haul networks is standard, single-mode (G.562-defined) fiber with dispersion in the 1,550-nm window, which limits the distance for OC-192 transmission. The majority of the legacy fiber plant cannot support high-bit-rate TDM. Vintage fiber has some attributes that lead to significant dispersion and would therefore be incompatible with high-bit-rate TDM.

State-of-the-art OC-192 SONET technology is a viable alternative for new fiber builds because the fiber parameters can be controlled through the placement of the appropriate fiber type and because new fiber-handling procedures can be accommodated. However, for existing fiber applications, the capability to install 10-Gbps TDM systems depends on fiber properties such as chromatic dispersion and PMD, which may differ from fiber span to fiber span. In addition, high-rate TDM systems require specialized fiber-handling and termination procedures when compared to lower-rate OC-48 systems. In contrast, current DWDM systems can transport up to 32 wavelengths at OC-48 each, giving an aggregate capacity of 80 Gbps on one fiber pair. This means WDM technology surpasses TDM in terms of the aggregate capacity offered on a fiber pair while maintaining the same fiber-handling procedures developed for OC-48 TDM systems.[13] SONET and WDM systems typically have built-in forward error correction (FEC) features that provide a 10^{-15} bit error rate (BER). Figure 5-25 depicts an example of typical interfaces supported by an OC-192 NE.

DWDM has a number of advantages over even the latest TDM option STM-64/OC-192. However, you must recognize that there are instances in which TDM may offer a better solution than DWDM. The recently introduced fiber NZDSF, for example, is flexible enough for the latest TDM equipment, but it is expensive and may limit the carriers' ability to migrate to the greater bandwidth available through DWDM at STM-16/OC-48 rates.

Figure 5-25 Example of interfaces.



Tributaries Supported on OC-192 Network Elements

- OC-12/12c 1310-nm Optics with 1+1 Protection
- STM-4/4c 1310-nm Optics with 1+1 Protection
- STS/OC-48/48c 1310/1557-nm Optics with 1+1/1:1 Protection
- STM-16/16c 1310/1557-nm Optics with 1+1/1:1 Protection

Many other service interfaces such as DS1/T1, DS3/T3, OC-3/3c, 10/100BaseT Ethernet, and MPEG2 digital video are also available in conjunction with subtending shelves.

5.2.4 Standardization Efforts

The wavelengths of the light output by laser diodes (LDs) have been chosen to minimize transmission loss across the fiber cable span and to utilize the highest sensitivity ranges of receivers. The most commonly used wavelengths employed by laser diodes are centered on 850, 1,310, and 1,550 nm. The ITU has recently standardized a grid of wavelengths for use in DWDM systems. This grid defines a minimum channel spacing of 100 GHz in the frequency domain corresponding to approximately 0.8 nm in the wavelength domain.[13]

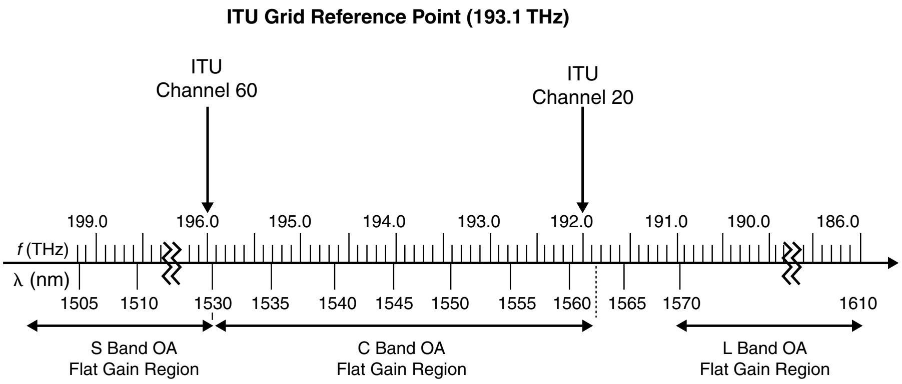
The ITU has defined in standard G.692 a channel optical frequency grid based on 100-GHz spacing from a reference frequency of 193.1 THz = 1,552.52 nm (see Figure 5-26). Figure 5-27 shows the grid from a different graphical perspective. A channel plan describes a number of factors that need to be taken into account in an optical communications network including the number of channels, channel spacing and width, and channel center wavelengths. Although this grid defines the standard, users are free to use wavelengths in arbitrary ways and choose from any part of the spectrum. Users can also deviate from the grid by extending the upper or lower bounds or by spacing the wavelengths more closely, typically at 50 GHz, to double the number of channels. Figure 5-28 depicts the usage of the frequency bands in 8-, 16-, 24-, or 32-lambda systems. However, interworking will suffer if vendors utilize variants of the basic standard.

Figure 5-26
ITU-T G.692 grid.

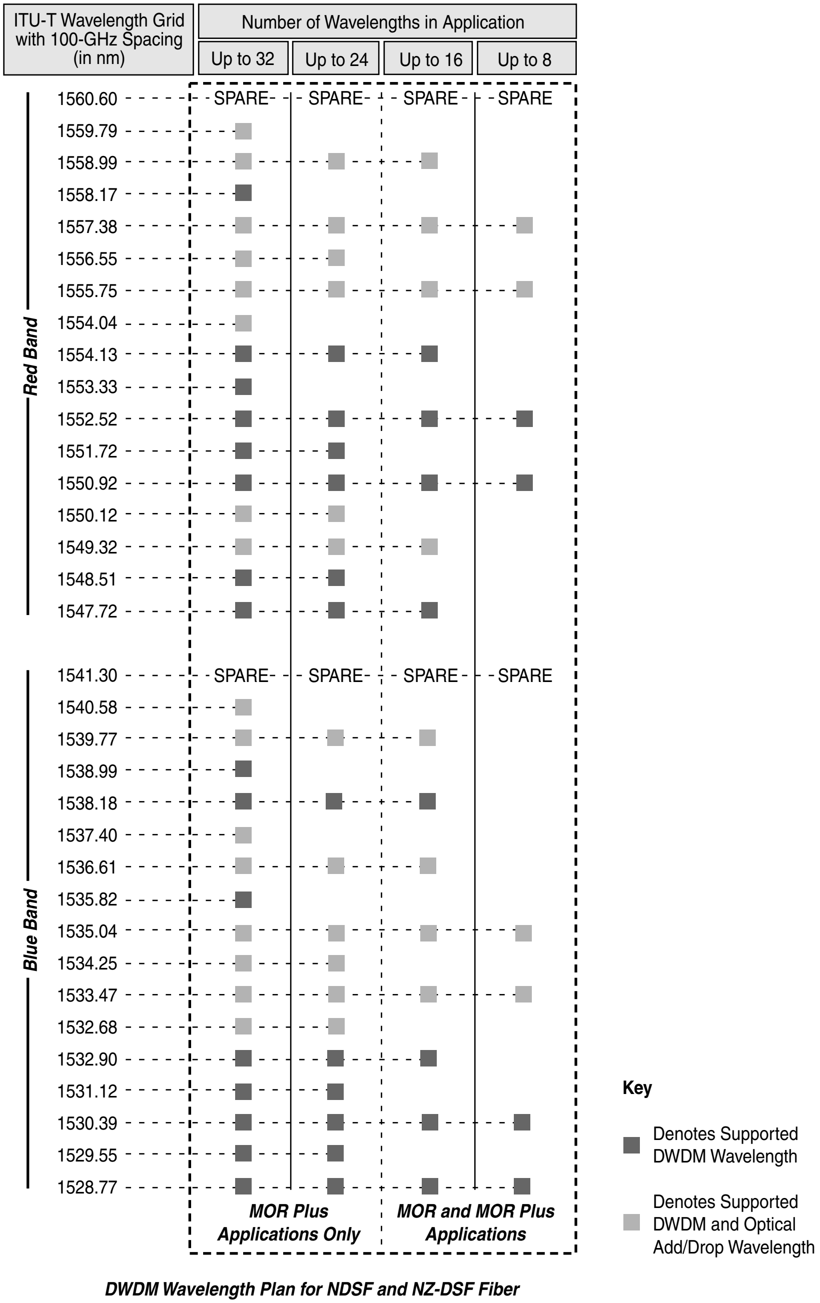
Frequency (THz[1])	Wavelength (nm[2])	Frequency (THz)	Wavelength (nm)	Frequency (THz)	Wavelength (nm)
196.1	1528.77	194.6	1540.56	193.1	1552.52
196.0	1529.55	194.5	1541.35	193.0	1553.33
195.9	1530.33	194.4	1542.14	192.9	1554.13
195.8	1531.12	194.3	1542.94	195.8	1554.94
195.7	1531.9	194.2	1543.73	192.7	1555.75
195.6	1532.68	194.1	1544.53	192.6	1556.56
195.5	1533.47	194.0	1545.32	195.5	1557.36
195.4	1534.25	193.9	1546.12	192.4	1558.17
195.3	1535.04	193.8	1546.92	192.3	1558.98
195.2	1535.82	193.7	1547.72	192.2	1559.79
195.1	1536.61	193.6	1548.51	192.1	1560.61
195.0	1537.40	193.5	1549.32	192.0	1561.42
194.9	1538.19	192.4	1550.12	191.9	1562.23
194.8	1538.98	193.3	1550.92	191.8	1563.05
194.7	1539.77	193.2	1551.72	191.7	1563.86

[1] THz = Terahertz
[2] nm = Nanometer

Figure 5-27
ITU-T grid on the electromagnetic spectrum.



DWDM systems with open interfaces give carriers the flexibility to support SONET/SDH, asynchronous protocols (for example, DS3), ATM, Frame Relay, and other protocols over the same fiber pair. Open systems also eliminate the need for additional high-performance optical transmitters to be added to a network when you need to interface with specific protocols (see Figure 5-29). Open WDM systems enable service providers to adapt new technologies to the network through the use of off-the-shelf, relatively



	ITU-T Wavelength Grid with 100-GHz Spacing (in nm)	Number of Wavelengths in Application: Up to 32	Up to 24	Up to 16	Up to 8
Red Band	1560.60	SPARE	SPARE	SPARE	SPARE
	1559.79	□			
	1558.99	□	□	□	
	1558.17	■			
	1557.38	□	□	□	□
	1556.55	□	□		
	1555.75	□	□	□	□
	1554.04	□			
	1554.13	■	■	■	
	1553.33	■			
	1552.52	■	■	■	■
	1551.72	■	■		
	1550.92	■	■	■	■
	1550.12	□	□		
	1549.32	□	□	□	
	1548.51	■	■		
	1547.72	■	■	■	
Blue Band	1541.30	SPARE	SPARE	SPARE	SPARE
	1540.58	□			
	1539.77	□	□	□	
	1538.99	■			
	1538.18	■	■	■	
	1537.40	□			
	1536.61	□	□	□	
	1535.82	■			
	1535.04	□	□	□	□
	1534.25	□	□		
	1533.47	□	□	□	□
	1532.68	□	□		
	1532.90	■	■	■	
	1531.12	■	■		
	1530.39	■	■	■	■
	1529.55	■	■		
	1528.77	■	■	■	■
		MOR Plus Applications Only		*MOR and MOR Plus Applications*	

DWDM Wavelength Plan for NDSF and NZ-DSF Fiber

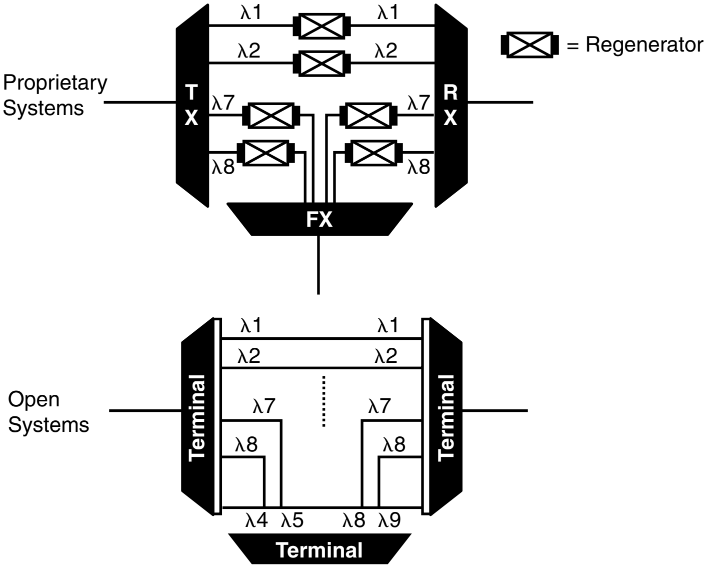
Key

■ Denotes Supported DWDM Wavelength

□ Denotes Supported DWDM and Optical Add/Drop Wavelength

Figure 5-28
Grid-point frequency usage in various (D)WDM systems.

Figure 5-29 Open versus proprietary systems.

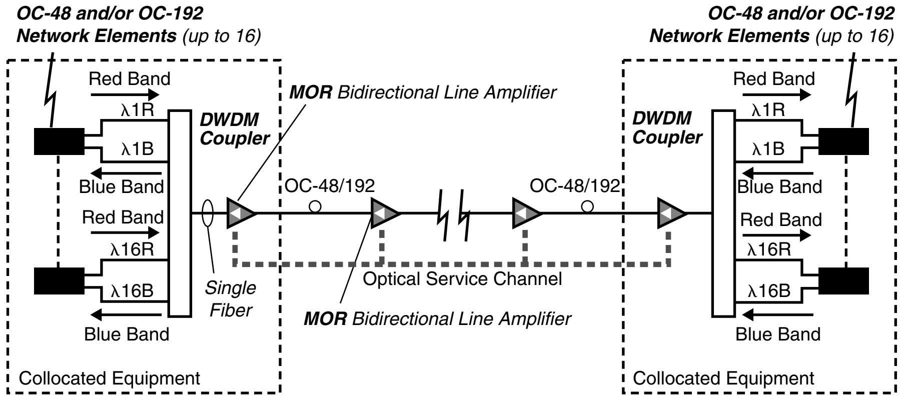


inexpensive, and readily available transmitters. In contrast to DWDM equipment based on proprietary specifications, systems with open interfaces provide operators greater freedom to provision services and reduce long-term costs. Proprietary systems, in which SONET/SDH equipment is integrated into the optical multiplexer/demultiplexer unit, are usable for simple point-to-point configurations; however, they require costly additional transmission equipment when deployed in meshed networks. DWDM systems that comply with the ITU channel plan reassure carriers that they can employ equipment from a variety of suppliers, although generally carriers do not mix equipment from two suppliers on the same ring because of network management (and more generally, Operations, Administration, Maintenance, and Provisioning [OAM&P]) considerations.

5.2.5 Example of High-End DWDM System

An example of a high-end DWDM system is Nortel's S/DMS Transport-Node. This system offers high-capacity transport systems operating at either 2.5 Gbps (OC-48) or 10 Gbps (OC-192). Figure 5-30 depicts a schematic view of the product. DWDM arrangements are supported in both OC-48 and OC-192 applications, permitting aggregate span capacities of up

Figure 5-30 Schematic view of Nortel products.



Notes:

λ1R/B thru λ16R/B denote individual wavelengths in red/blue band.

Multiwavelength Optical repeater (MOR) for extended reach in multiwavelength DWDM application up to 16 wavelengths.

Up to 32 wavelengths supported using MOR optical amplifiers.

Red band = 1547.5 to 1561 nm, blue band = 1527.5 to 1542.5 nm.

to 320 Gbps in a fully expanded configuration and employing the OC-192 line rate and 32 wavelengths. Hybrid configurations are available as well, enabling both OC-48 and OC-192 systems to be combined in the same DWDM application. Concurrent support for mixed SONET and SDH tributaries from the same OC-192 network element yields lower life-cycle costs and also offers many operational efficiencies. Included in S/DMS TransportNode's high-capacity transport family are the multiwavelength optical repeaters (MOR and MOR Plus), which are next-generation optical amplifiers that inexpensively extend system reach between widely separated service-terminating or regenerator sites on fiber routes. S/DMS TransportNode encompasses both the SONET/SDH and optical layers.

5.3 DWDM Advocacy for Metro Access/Metro Core

5.3.1 Issues

The following observations from a key vendor are refreshing: "Although capacity per fiber is important in the local networks, the value equation is different because the topologies and traffic patterns are more meshed, resulting in shorter spans with less capacity per span. The capacity demand is still within the realm of TDM technology and the cost advantages of optical amplification cannot be fully exploited because of the shorter distances. As such, WDM technology must bring a richer set of values to meet the diverse challenges of the local network."[13]

In long-haul networks, the major portion of the network transport cost is due to the numerous regenerators along a link (typically every 35 to 70 miles). Normally, not only do you need to do the amplification, but you also need an O-E conversion, followed by an E-O conversion (resulting in a full O-E-O conversion). Utilizing OLAs reduces this cost by eliminating the O-E-O conversion. As the cost of optical amplifier technology drops, WDM becomes increasingly more economical in long-haul networks. However, in local networks, where distances are shorter, the bulk of the *core* network transport cost is due to the multiplexers and crossconnects. In the *access* network, it is the cost of the fiber spur. Hence, reductions in the cost of these components will prove advantageous to the bottom line. The one major factor driving the potential success of DWDM in metro access/metro core is the cost per unit (specifically per lambda) per building. The cost per lambda should be in the range of $2,000 to $10,000; the cost per building should be in the range of $10,000 to $20,000; and the system should carry 4, 8, or 16 protected lambdas (8, 16, or 32 distinct lambdas).

It is important to consider how many broadband customers actually reside in one building. The model of Chapter 1 indicated the results of Table 5-4.[16] It is worth noting that a carrier that needs to connect approximately 300 buildings using 36 metro access rings of 8 buildings each and can pick up 5 to 10 broadband customers per building has a financial advantage over a follow-on carrier that needs to connect the same 300 buildings with an overlay of 36 new rings and only pick up 1 or 2 broadband customers per building. If the building had 3 broadband customers, the economics begin to work, but only when the per-building nodal cost is low (for example, in the case of mid-range, next-generation SONET rather than more expensive DWDM).

These results, grounded in a level of pragmatism, should prove to be instructive. Even a successful carrier will only have a handful of broadband customers in a building.[17] Hence, the deployment of high-channel-capacity WDM systems in the metro access is highly ineffective from a financial standpoint and an overkill. Even if you assume that there are eight buildings on a ring and each building has three broadband customers that a carrier is able to bring online, it needs to be noted that the typical wide area network (WAN) speed today is 10 to 100 Mbps. This equates to a maximum

Table 5-4

Broadband Customers per Building in the United States

	2002	2003	2004	2005	2006
Fibered buildings (CIR et al. stats) per city (top 30 U.S. cities) (*)	300	347	400	462	534
Broadband customer/building	5	7	10	15(#)	20(#)
Broadband customer/building, first mover (50%)	3	4	5	8(#)	10(#)
Broadband customer/building, second mover (25%)	1	2	3	4(#)	5(#)
Broadband customer/building, all other (25%)	1	1	2	3(#)	5(#)

(*) The remaining 60 cities have the other 50 percent balance, as modeled in Chapter 1.

(#) This number may in fact turn out to be high, but we follow the model's formulation.

of 2.4-Gbps per ring, which fits the profile of a next-generation SONET solution very well. If by 2006 the bandwidth need grows fourfold, OC-192 next-generation SONET solutions will still be adequate. However, again, it is important to understand that the issue is less technical and more financial. If a DWDM solution were to enjoy cost parity with a next-generation SONET solution, the planner would have no problem selecting the former over the latter.

In summary, the issues in metro optical networks are significantly different than in the long haul. Whereas long-haul networks are focused on lowering the cost per bit for point-to-point transport, metro networks are more concerned with cost-effective connectivity to a multitude of endpoints (that is, connectivity to a large set of buildings), service transparency, end-user protocol support, speed of provisioning, and end-to-end connectivity. During the 1990s, the metro network was primarily concerned with the transport of voice-based TDM circuits (DS1 and DS3) for private-line services. Recently, the issues related to the metro network are how to support the emergence of data services based on IP and lambdas, as well as how to support the continued growth of legacy TDM. Another related issue is how to build the metro network to support the shift from legacy TDM to a data-centric network based on IP.

Metro access and core networks have to carry many different types of traffic over an optical channel, regardless of the protocol (Ethernet, ATM, SONET, and so on) or bit rate (10 Mbps, 2.5 Gbps, and so on). Figure 5-31

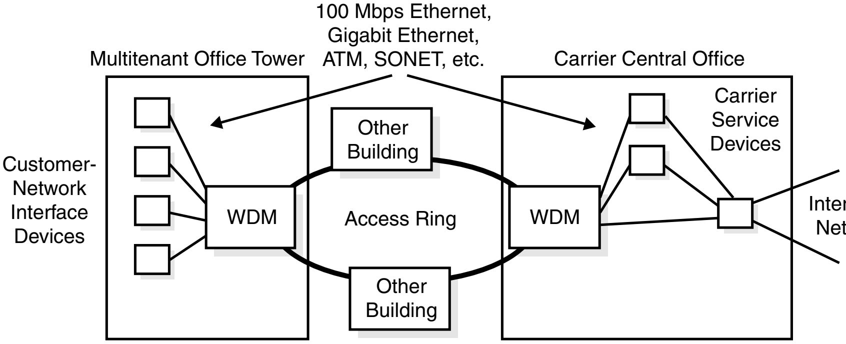


Figure 5-31 Metro DWDM applications example.

typifies an access network with protocol and bit-rate independence. Native data interfaces, such as Ethernet at various speeds, can be connected directly to the transport platform without additional adaptation. As we discussed in the Chapter 4, “Next-Generation SONET and Optical Architectures,” next-generation SONET targets precisely this segment. For Figure 5-31 to be more than just a picture, the cost per building for the DWDM needs to be $10,000 to $20,000 for GbE and 10 GbE applications and $5,000 to $10,000 for 10-Mbps and 100-Mbps transparent LAN service (TLS).

Nearly all enterprise data traffic now originates in Ethernet format (10 Mbps/100 Mbps/GbE), while the backbone is built on optical rates of 2.5 and 10 Gbps with STS-1 granularity. This is why it is important for metro WDM devices to support the data-oriented interfaces; this would not be a requirement for long-haul DWDM. On the other hand, a traditional (telco) metro optical network continues to focus primarily on legacy DS1 and DS3 granularity within the metropolitan area network (MAN) and through the backbone. A full-featured metro WDM must support both the traditional interfaces as well as the new data interfaces. Enterprise routers currently support direct OC-3c, OC-12c, and OC-48c/192c WAN links. These private-line interfaces on end-user equipment (that is, on intranet routers) are widely deployed because of their ubiquitous deployment throughout metro optical networks (for example, the extensive deployment of SONET transport and crossconnects). However, this comes at a cost. High-end SONET plug-ins can cost up to $50,000. Therefore, it is desirable to include these interfaces in the NE because of the economies of scale that can be secured by network providers, which can then be passed onto the enterprise user.

Equipment vendors working on WDM are realizing that data-oriented services are now consuming more network capacity than voice due to an increasing number of users and escalating capacity demands from each user. The protocols, bit rates, and interfaces used by the enterprise networks that support these applications are different from those used in the carrier networks. The former are increasingly Ethernet based; the latter are typically OC-*x* based. Conversion between the two domains is needed. The user can affect that conversion, but it has to be done at multiple places where the enterprise network touches the carrier's network or, better yet, it can be done by the carrier on behalf of the user. One example may suffice. Carriers have come to realize that it is better to carry voice over distances in digital format. However, the network does not impose that requirement to the user. All residential users generate analog voice over their telephone handsets and the network adapts the user's signal into a digital signal that is better suited for network handling. Similarly, because all enterprise networks are now based on Ethernet/IP/TCP technology, it would be preferable for the carrier's network to accept this handoff and then convert it, if necessary, to a SONET/WDM format. Because Ethernet interfaces are very cost effective for the end user compared to traditional WAN interfaces, TLS-based services (supported by Ethernet-ready edge NEs) would be an attractive offer to corporations. It also saves costs for the network provider by eliminating the adaptation functions to ATM. Although ATM can support speeds of tens of megabits per second, it does not support interfaces operating at 1 or 10 Gbps equally well. Native data interfaces for other protocols on edge NEs, such as ESCON or FICON, would provide similar values depending on user-specific requirements.

At the end of the day, however, native data interfaces are not necessarily intrinsic to WDM only. As noted, next-generation SONET aims at supporting these same interfaces. Hence, the choice between WDM and next-generation SONET is really a traffic-engineering consideration once we get past the initial cost. This raises the following question: How much growth can be anticipated in a metro access subnetwork?

By deploying a WDM-based photonic network, the service provider gains an access or interoffice transport infrastructure that is flexible and scalable. However, planners need to ascertain that what could turn out to be a high NE cost does not become multiplicative and all 8, 16, or 32 buildings on the ring are each forced to incur a high nodal cost. Although a long-haul network may be a two-point link, or at most be a ring/mesh with a half a dozen or so high-capacity points (say, in several states), metro access rings typically have a larger number of nodes on them, unless a nonredundant building access is utilized, as shown in Figures 1-8 through 1-12 in Chapter 1. A 16-building access ring using next-generation SONET with a

$30,000 device would cost the provider $480,000 in nodal equipment; a ring using DWDM with a $60,000 device would cost the provider $960,000.

There are also certain key OAM&P implications for metro core WDM networks. At this time, long-haul applications do not require dynamic management of individual wavelengths (except with new architectures such as Generalized MPLS (GMPLS), which is discussed in Chapter 8, "GMPLS in Optical Networks"). Therefore, one optical surveillance channel carried on a dedicated wavelength is sufficient to control and monitor all remote WDM network elements such as OLAs. Photonic networks in metropolitan access and interoffice applications manage individual wavelengths. Because individual wavelengths have different endpoints and take different paths across the network, one optical surveillance channel carried on a separate wavelength is not sufficient. Fault information is required for and must accompany each wavelength. Per-wavelength fault information is needed for fault detection and isolation, which are necessary to support survivability mechanisms and trigger fast maintenance and repair actions. In photonic networks, service providers can also verify the connectivity and monitor the performance of each connection.[13]

Proponents make the case that with metro DWDM, user requests for increased bandwidth or different protocols can be filled quickly, so the network provider can realize increased revenue. This may be true in the metro core network, but it is unlikely that it can be achieved in general in the metro access. New services such as optical-channel leased lines that provide end-to-end protocol and bit-rate-independent connections can be offered, affording new revenue for the network provider. In WDM-based photonic networks, transport interfaces that are specific to protocol and bit rates are no longer required, minimizing the network provider's inventory and operating costs.

On the other hand, it is important to keep in mind that the market for traditional handoffs from intranet routers will continue to exist for many years and may in fact exceed the Ethernet-handoff market for the foreseeable future, as anecdotal information[18] tends to show. The camp that supports native data interfaces asks the following rhetorical question: *Given that WAN traffic from the enterprise into the metro network is primarily Ethernet data today and it is likely to continue to dominate in the future, should the conventional view that legacy private line services are the most efficient means for metro optical networking be challenged?* In the long term (three to five years out), metro optical networking based on Ethernet native LAN interfaces is likely to become prevalent. This is because there is additional cost for Ethernet router vendors to adapt native data interfaces into private-line interfaces (DS1, DS3, and OC-*x*) using what are referred to as *telecom adapters*. The additional cost translates into private-line interfaces costing as much as three to four times more than native LAN interfaces (100Base-T and GbE). In addition, the cost of GbE interfaces is expected to drop rapidly as their deployment rises in enterprise networks. Bandwidth growth has been limited to private-line-like bandwidth increments such as DS1, DS3, or OC-*x*. Because these private-line interfaces are provisioned as physical connections from the customer premises onto the metro network, incremental bandwidth needs require new physical connections to be established. However, as noted, the reality is that the legacy interfaces will be around for a number of years. Clearly, most carriers have invested billions of dollars in SONET optical products, ATM switches, and digital crossconnect system (DCS) crossconnects; hence, they are highly motivated to price and design these services right in order to achieve a return on their investment, particularly in the new telecom economics market. Eventually, carriers will migrate to new technologies. The challenge is making a graceful migration from the embedded network to a network based on data-oriented interfaces.

5.3.2 Network Management in DWDM

A critical yet often understudied part of any telecommunications network is the cost of OAM&P. Part of this issue drills down to the management system. Dependable and easily accessible network management services can increasingly become a distinguishing characteristic of high-performance, high-capacity systems. Well-designed DWDM systems include integrated, network management programs that are designed to work in conjunction with other operations support systems (OSSs) and are compliant with the standards the ITU has established for the telecommunications management network (TMN). Simple Network Management Protocol (SNMP) capabilities may also be useful, particularly for new carriers that are more data focused and may evaluate a DWDM in contrast with GbE, 10 GbE, or Resilient Packet Ring (RPR) systems developed by the "Ethernet people" that will support SNMP. Systems can utilize an optical service channel that is independent of the working channels of the DWDM product to create a standards-based data communications network that enables service providers to remotely monitor and control system performance and use. This network manager communicates with each node in the system and also provides dual-homing access and self-healing routing information in the event of a network disruption.

5.4 Example of Optimized Edge Equipment

This section describes an example of what kind of edge technology may be of practical interest to prospective metro-focused carriers. The key point is that DWDM plays some role, but a multivalued optimized footprint/price-point for edge equipment is critical. Hence, a blind formula that states "deploy DWDM for metro applications" leads to severe financial distortion for prospective carriers and is a major difficulty in deploying a sustainable, short-EBITDA-positive enterprise. Successful carriers realize that they need a multiarchitecture, multitechnology, multiservice, multi-QoS, multimedia, secure, reliable, and affordable service-delivery platform. These types of products are being planned by EtherCarrier Communications (which is located in Red Bank, New Jersey).

EtherCarrier is a newly launched developer of price-optimized, carrier-class, optical, next-generation edge and central office (CO) devices for GbE and 10 GbE in the WAN technology. It is to be used by next-generation information carriers including Ethernet Local Exchange Carriers (ELECs), metropolitan area service providers (MASPs), Fiber LECs (FLECs), Competitive LECs (CLECs), Incumbent LECs (ILECs), Internet service providers (ISPs), and high-end corporate customers with campuses, privately owned fiber, or leased fiber. The developer takes advantage of conspicuous secular trends in the move toward full-speed, end-to-end, 1 and 10 GbE in the WAN to support native Ethernet and IP over optics communication networks that new carriers will set out to deploy in the next couple of years—moving on from the ashes of the CLEC maturation, saturation, and debacle that was seen at the dawn of 2000. The goal of the developer is to supply carrier-class, cost-effective, fiber-optic network edge devices to the carrier industry to meet the ever-increasing market demands to deliver converged broadband IP-based data/Internet/internetworking services over gigabit infrastructures that make use of the latest optical and customer-control technologies. These devices fill a market void.

The company develops optical network equipment for both edge and CO applications based on the latest optics, GbE, and 10 GbE WAN technologies. This kind of technology is being referred to as *third-generation* (3G) optics. Optical edge equipment will be deployed in customer buildings and connected to the CO optical equipment over a traditional fiber facility. The edge equipment enables fiber-optic pair gain, where multiple customers in a building are each provided between 1 and 10 Gbps, but all utilize the same (protected) access fiber pair. The goal is to extend the advances seen in the optical space over the past few years in the long-haul segment to the critical metro access and metro core segments of the market where there is a void and a major continuing opportunity. The products are not principally focused on increasing raw bandwidth (although that may be achieved as a consequence), but are focused on achieving cost-effective connectivity instead, which continues to be a real need in the industry.

The developer's flagship set of metro edge products (specifically, advanced OADMs) provides low-cost, right-sized access vehicles, including vehicles based on WDM and CWDM. The goal of the products is to support ubiquitous broadband line-rate *Ethernet Connectivity Anywhere*™ data/Internet service. With Ethernet Connectivity Anywhere, the customer has access to an end-to-end IP-ready connectivity/internetworking service over a more native interface that is tailored to his or her needs, and the customer is able to secure Ethernet-based access speeds, rather than the traditional lower speeds (such as 1.5 Mbps). The goal is to develop a service-optimized metro edge suite of smart access systems over advanced optics. This fiber-multiplication, multicustomer-per-edge, optical, metro core, carrier-class technology enables service providers to economically relieve the bandwidth scarcity and connectivity bottlenecks that now exist in the access segment of the public service backbone.

This legacy bottleneck environment is partially due to entrenched costs of metro area networking. Today you would need a carrier to deploy a building-resident mux that could cost $40,000, $80,000, or even $120,000. It follows that the monthly recurring charge (MRC) to the customer has to be high. This area is now beginning to change with the widespread adoption of technologies such as GbE and WDM. However, DWDM remains relatively expensive and is best suited for long-haul rather than access applications. EtherCarrier's products are targeted to access applications, where technologies such as CWDM may be better suited. Using metro DWDM, a service provider can provide OC-48 (the equivalent of 55 DS3s) to a customer premise for a cost to the carrier of about $50,000 per month (in order for the carrier to be able to recover costs). Using CWDM, this service can be provided for about $20,000 per month, or less. An important difference between metro core and long-haul systems is making DWDM an option on the system, not a requirement. Cards that support DWDM will significantly increase the cost of the system (because each channel requires its own components), yet there are situations where DWDM is not required. Such a situation would arise when a service provider has to pull new fiber to a location, but, for a period of time, would only need to provision one wavelength. Here, single-wavelength SONET would suffice, and CWDM or DWDM could be added later as bandwidth requirements expanded. Figure 5-32 identifies the opportunity space for these types of metro products.

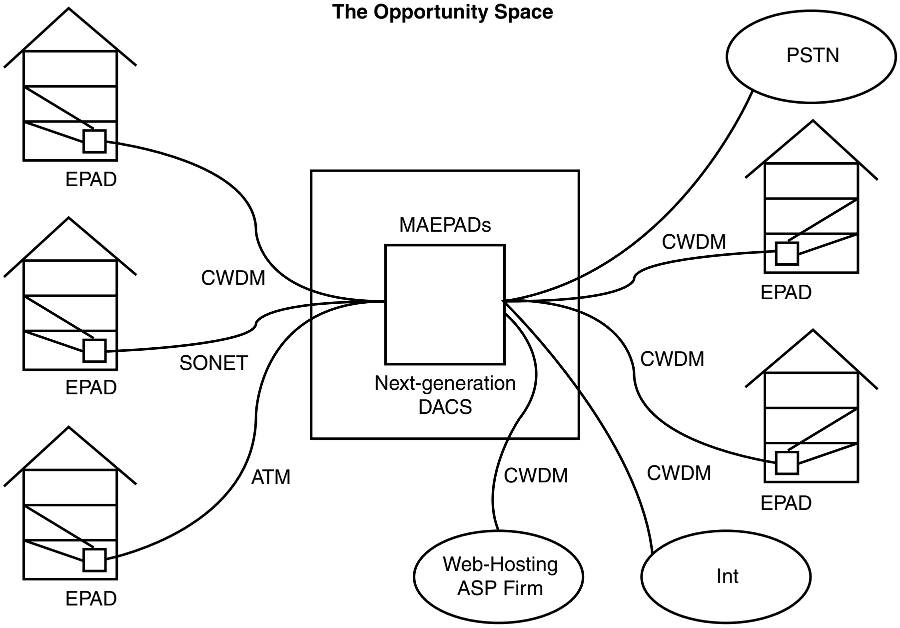


Figure 5-32 Opportunity space for price-optimized edge devices.

The product set includes the following:

- Edge Packet Aggregation Devices (EPADs) (see Figures 5-33 through 5-37)
- CO Massively Arrayed EPADs (MAEPADs) of various footprints (see Figure 5-38)

Specifically, the developer plans to manufacture a family of *edge products*, ranging in price from $2,500 to $40,000, which support a variety of access, network, and routing protocols; as well as *CO products*, ranging in port count from 32 to 128, which also support various speeds, protocols, and media.[19] The product space for the EPADs is shown in Table 5-5, which shows each product optimized to a class of carrier applications and a cost-benefit pencil.

The protocol model of the modular EPAD is shown in Figure 5-39. Point A depicts a routing or bridging function (when null). Point B depicts the network-side data link layer (DLL) technology. The maximum-form EPAD product is shown in Figure 5-40; however, different subproducts are developed to be optimal for the specific application at hand.

Figures 5-41 through 5-46 illustrate the feature/cost profile addressed. The space shown depicts an increasing customer- and network-side capability for the edge device, along with the cost range (from $2,500 to $40,000).

Table 5-5

Product Scope

Class	Subclass	Description
EPAD-100		
	EPAD-100/IPA	Entry-level edge router with IP over ATM
	EPAD-100/IP0	Entry-level edge router with IP over link
	EPAD-100/0A	Entry-level edge bridge over ATM
	EPAD-100/00	Entry-level edge bridge over link
EPAD-200		
	EPAD-200/IPE	Mid-level edge router with IP over WAN Ethernet
	EPAD-200/IPA	Mid-level edge router with IP over ATM
	EPAD-200/IP0	Mid-level edge router with IP over link
	EPAD-200/0E	Mid-level edge bridge over WAN Ethernet
	EPAD-200/0A	Mid-level edge bridge over ATM
	EPAD-200/00	Mid-level edge bridge over link
EPAD-300		
	EPAD-300/IPE	Higher-level (optical) edge router with IP over WAN Ethernet
	EPAD-300/IPA	Higher-level (optical) edge router with IP over ATM
	EPAD-300/IP0	Higher-level (optical) edge router with IP over link
	EPAD-300/0E	Higher-level (optical) edge bridge over WAN Ethernet
	EPAD-300/0A	Higher-level (optical) edge bridge over ATM
	EPAD-300/00	Higher-level (optical) edge bridge over link
EPAD-400		
	EPAD-400/IPE	High-end optical edge router with IP over WAN Ethernet
	EPAD-400/IP0	High-end optical edge router with IP over link
	EPAD-400/0E	High-end optical edge bridge over WAN Ethernet
	EPAD-400/00	High-end optical edge bridge over link
EPAD-500		
	EPAD-500/IPE	High-end low-cost optical edge router with IP over WAN Ethernet
	EPAD-500/IP0	High-end low-cost optical edge router with IP over link
	EPAD-500/0E	High-end low-cost optical edge bridge over WAN Ethernet
	EPAD-500/00	High-end low-cost optical edge bridge over link

Figure 5-33
EPAD-100.

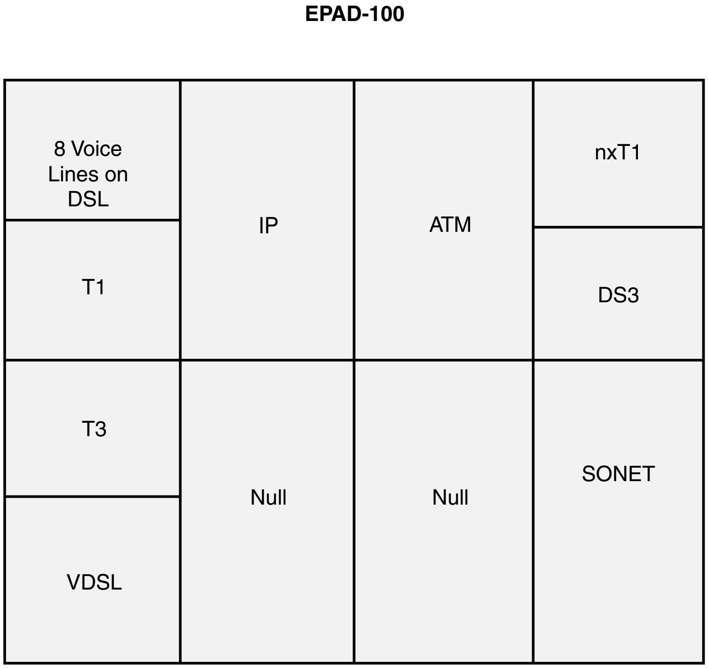


Figure 5-34
EPAD-200.

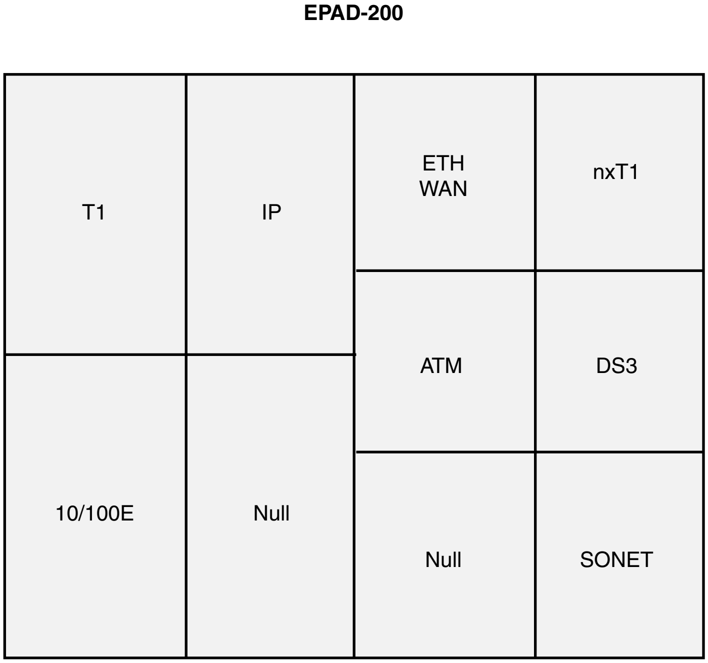


Figure 5-35 EPAD-300.

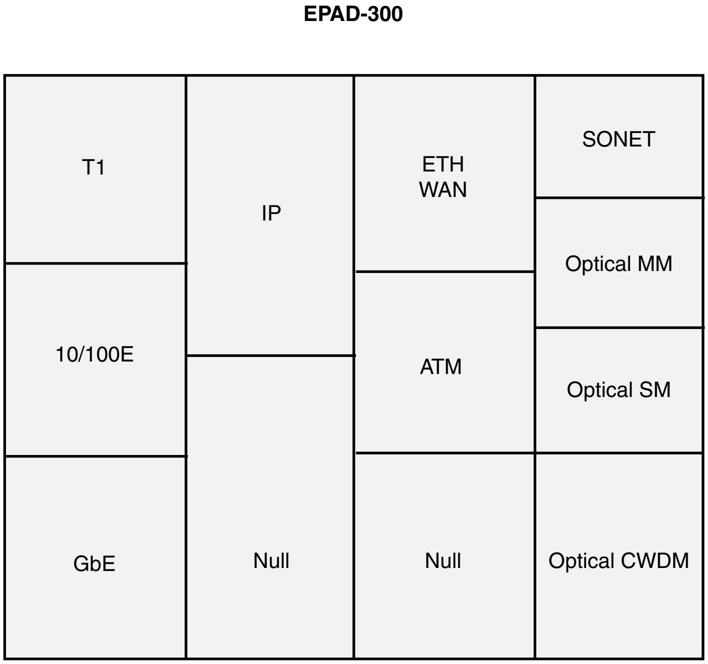


Figure 5-36 EPAD-400.

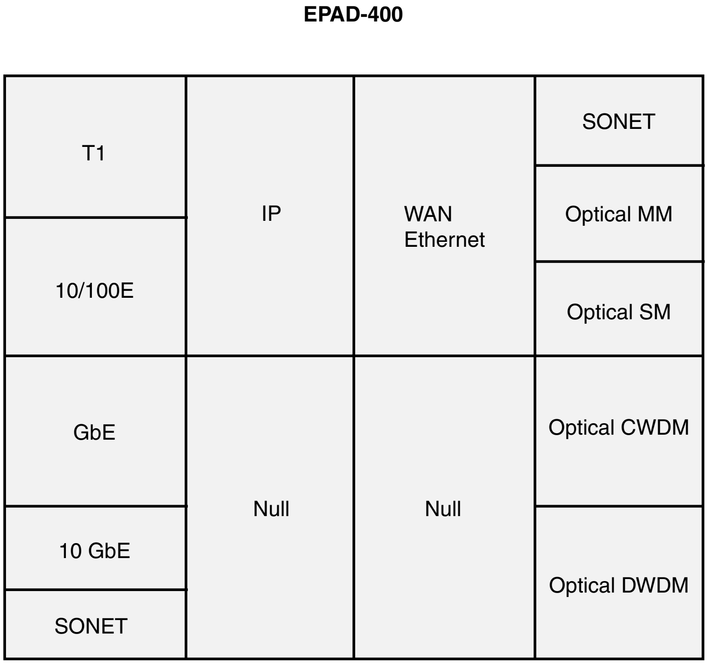


Figure 5-37
EPAD-500.

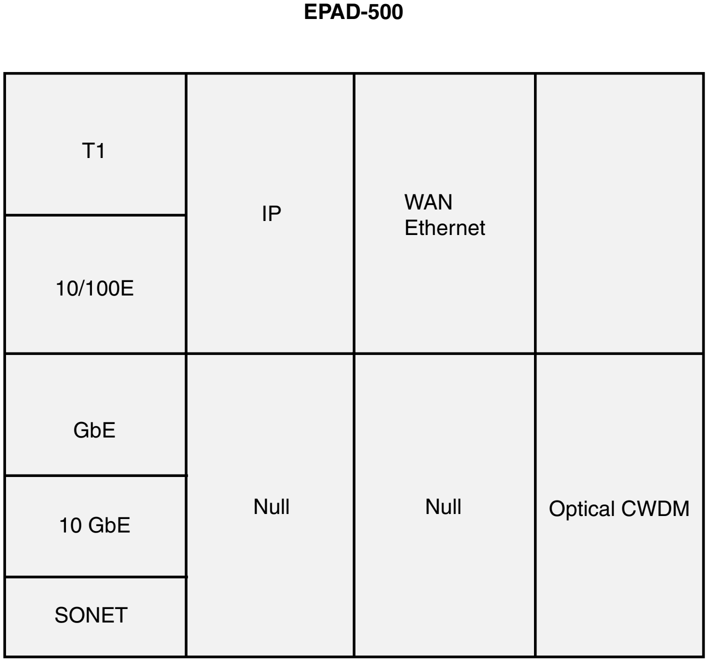
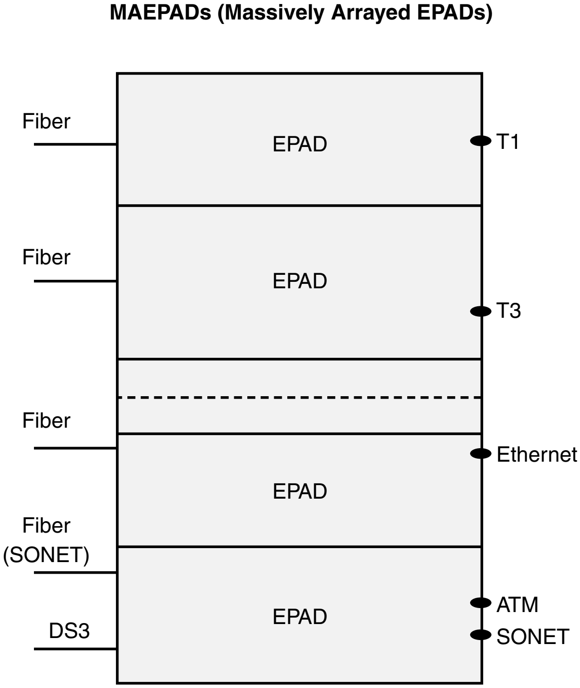


Figure 5-38
MAEPADs.



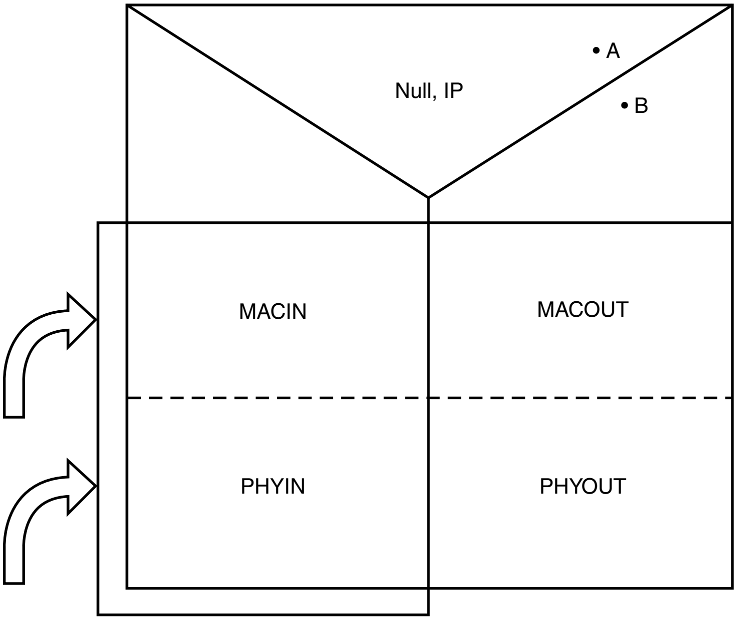


Figure 5-39
Protocol model for price-optimized edge devices.

In each figure (5-33 to 5-38), a number of products are indicated, which are each targeted to a specific carrier application. The value proposition is to develop edge devices that meet a certain functional/price-point envelope, as shown by Figures 5-33 to 5-38.

Note that the systems in Figure 5-33 have no Layer 2 or Layer 3 capabilities. The systems in Figure 5-34 have a Layer 2 capability over ATM, but no Layer 3 capability. The systems in Figure 5-35 have a Layer 2 capability over a WAN Ethernet (10 GbE WAN interface sublayer [WIS]), but no Layer 3 capability. The systems in Figure 5-36 have no Layer 2, but do have Layer 3 capability. Figure 5-37 systems have an ATM capability at Layer 2 and Layer 3 capability as well. The systems in Figure 5-38 have a WAN Ethernet (10 GbE WIS) capability at Layer 2 and also a Layer 3 capability.

At the component level, there are really no shortcuts that lead to near-term cost reduction other than the component suppliers continuing to improve yields, increase volumes, and reduce manufacturing costs.[20] There are, however, specific component and module strategies coming out of the component community that address metro needs for reduced costs in specific areas. You should focus less on cost reductions at the component level (via some upcoming technical breakthrough—however, see the following discussion) and more on getting the right-sized finished product in terms of

Figure 5-40 Protocol combination space.

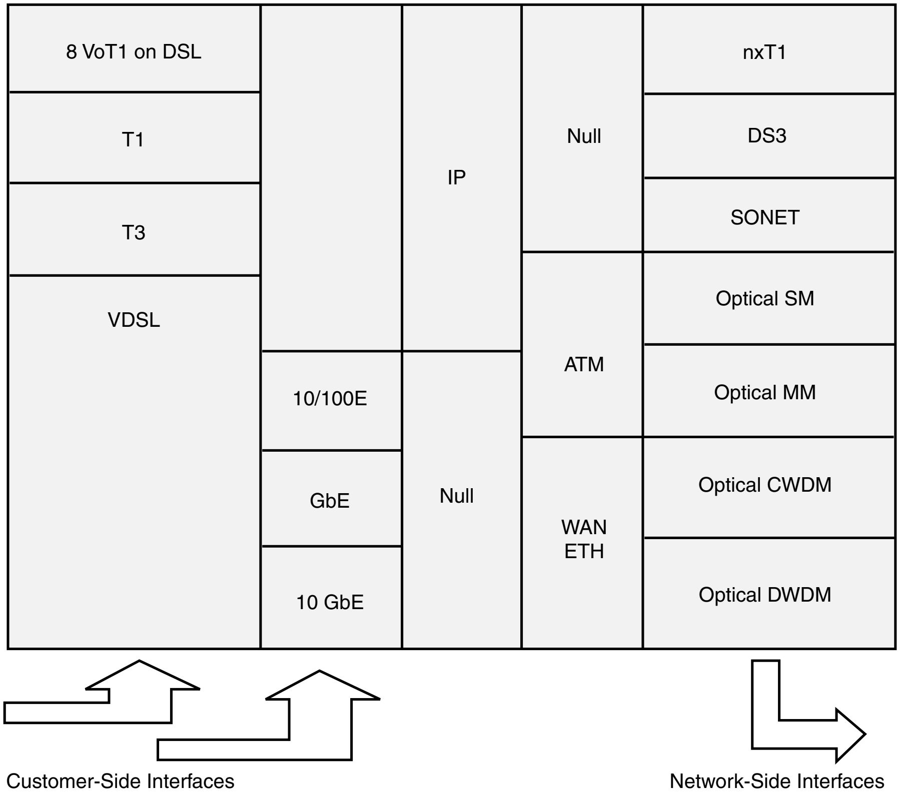
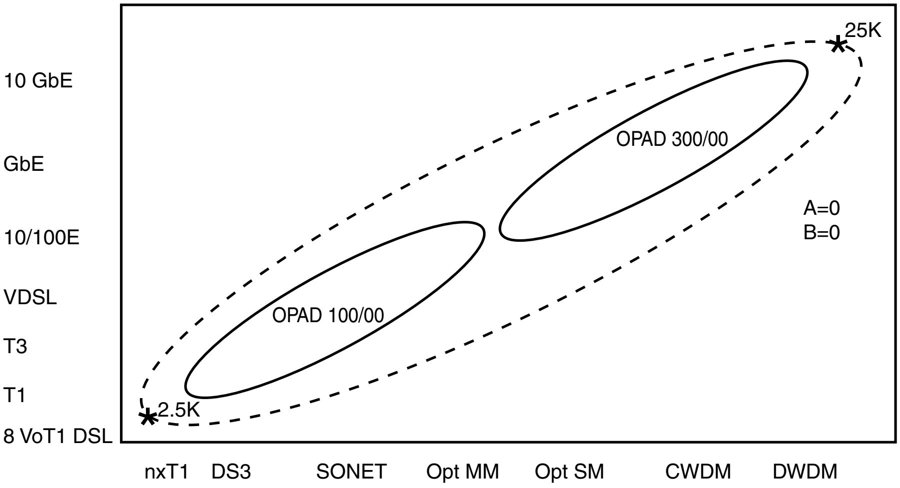


Figure 5-41 Opportunity space for price-optimized edge devices.



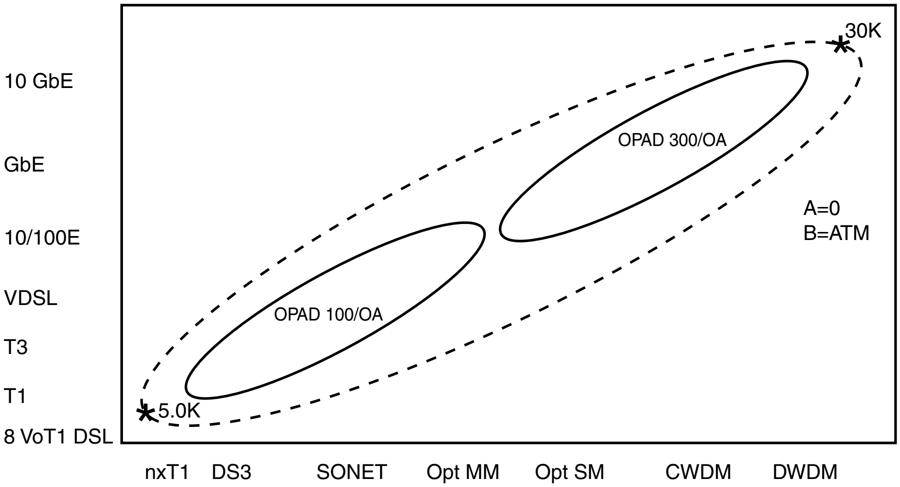


Figure 5-42 Opportunity space for price-optimized edge devices—another case.

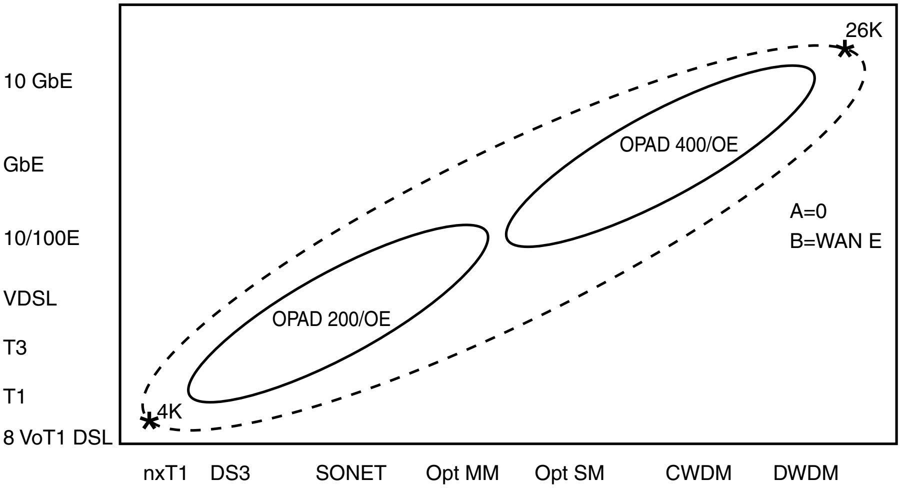


Figure 5-43 Opportunity space for price-optimized edge devices—another case.

port density, power budget, maximum capacity throughput, and handoffs. For example, in nearly all instances using a DWDM product to support a typical commercial building, it is an overkill, way too expensive, and leaves a lot of unused capacity in the systems.

Some of the possible technologies that could lower edge-device/transmission system costs include, but are not limited to, the following:

- **CWDM** In the component industry, a significant amount of effort is focused on reducing channel spacing to enable DWDM systems to

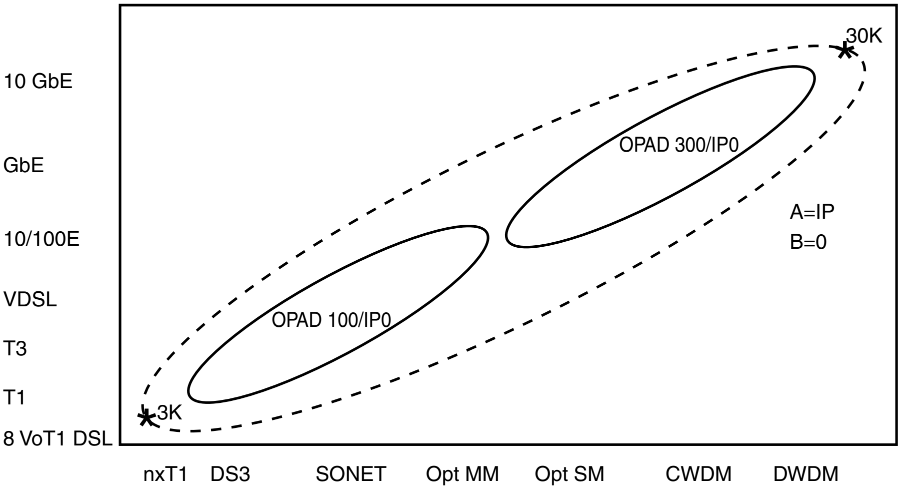


Figure 5-44 Opportunity space for price-optimized edge devices—another case.

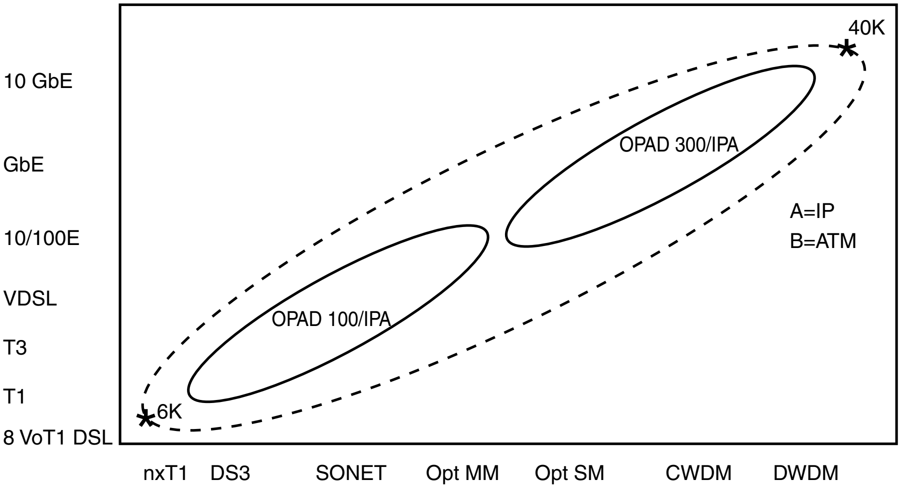
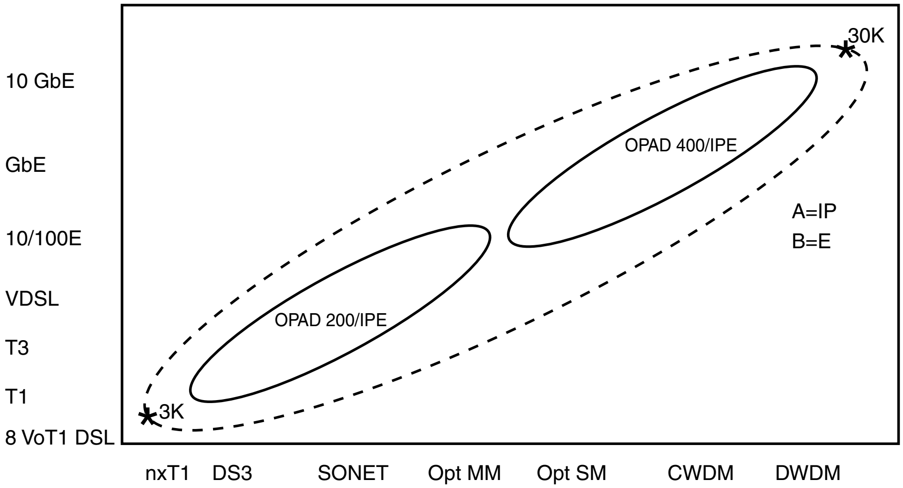


Figure 5-45 Opportunity space for price-optimized edge devices—another case.

support more wavelengths. Today, typical wavelength spacings using thin-film filters are at 200 GHz, moving to 100 GHz and even 50 or 25 GHz. The precision associated with this channel spacing results in challenging yields and higher component costs. However, the metro market, especially metro access, will operate at much lower channel density, allowing as much as several nanometers (nm) between channels (that is, a 4- or 8-channel system could operate in the neighborhood of 1,310, 1,480, 1,540, and 1,560 nm). Meanwhile, the

Figure 5-46 Opportunity space for price-optimized edge devices—another case.



cost of producing less-precise components (such as lasers, filters, and so on) has fallen as yields have improved. This topic is covered in more detail in Chapter 6, "Coarse Wavelength Division Multiplexing (CWDM) Technology."

- **Vertical Cavity Surface-Emitting Laser (VCSEL)** With current technology, lasers used in networking systems emit the signal from the edge of the diode. A substantial amount of research is currently under way to develop lasers that emit the signal from the surface of the diode. VCSELs have been around for quite a number of years, but engineering and cost improvements are still being sought. The main feature of VCSELs that makes them potentially cheaper than edge-emitting lasers is that an entire wafer can be tested without dicing it into individual diodes. Currently, edge-emitting lasers for communication applications are expensive, around $1,000 each. In addition, the packaging process with VCSELs is less complicated and therefore less costly. VCSELs have the potential to cost much less, provided that they can be produced at the appropriate communication wavelengths.

  Currently, most VCSELs operate at 840- and 980-nm wavelengths. The cost of the individual 840-nm VCSEL is in the $2 to $20 range. These VCSELs are used in data communication applications for intracomputer and intercomputer communication systems that have distances in the 10-m range. VCSELs in the 980-nm range are not currently used because their emission straggles several wavelengths, making transmission problematic.

The key issue VCSEL developers are grappling with is the output power that a VCSEL emits at a single wavelength, the so-called single-mode VCSEL. The output power of these VCSELs is in the 1- to 4-milliwatt range, which is far below that of edge-emitting lasers (edge-emitting lasers currently provide power in the hundreds of milliwatts). Clearly, the gap between VCSEL and edge-emitting products is large, but the power requirements in metro applications are considerably lower than long-haul applications because the distances are much shorter. Ten milliwatts of single-mode output power (double the current performance of VCSELs) may be adequate for 40-km rings, which are typical of metro applications. If these power outputs can be achieved, then VCSEL technology will likely be a major cost-reduction factor for optical-networking systems in the metro market.

- **Sub-$1,000 amplifiers** Because metro optical systems operate over much shorter distances than long haul, one basic approach of system designers and network planners is to use as little amplification as possible. Perhaps some geographically spread-out metro areas (such as Atlanta, San Jose, and so on) could need some amplification. However, additional capabilities such as wavelength switching and routing will increase insertion loss and other types of signal degradation in the network. Although large repeater amplifiers, priced at about $20,000, continue to grow in complexity and power, a new class of amplifiers is emerging. These amplifiers are simple and typically have only one pump diode. The major application is overcoming losses due to the proliferation of optical-networking elements, such as OADMs, OXCs, and other equipment, which results in a loss of optical power.

  There are two major types of metro amplifier, the semiconductor-optical amplifier (SOA) and EDFA. SOAs have an advantage in that they exhibit gain across a much wider wavelength spectrum, specifically over 300 nm versus a less than 100-nm spectrum for EDFAs. However, SOAs have a limitation in that they cannot amplify multiple wavelengths simultaneously without crosstalk between the channels. SOAs will likely find a niche in applications where the wavelengths to be amplified are outside the EDFA gain region and also as optical switch devices. Assuming that there is continued cost reduction, EDFAs should dominate the market for low-cost amplifiers in the 1,500- to 1,600-nm wavelength range. EDFAs targeted for the metro market achieve their cost reduction through a scaled-down component architecture, innovative design, and automation. There are already amplifiers that cost about $3,000 targeted at the metro market, and it is possible that in the future amplifier costs will approach the sub-$1,000 price range.

- **Direct router access to optics** Capabilities related to Optical Domain Service Interconnect (ODSI) are emerging. The means by which electrically based network devices such as IP routers, Layer 2 switches, and metro transport devices interoperate with an optical core deserves attention. Proponents advance the idea that the optical layer must effectively internetwork with the devices listed, bridging the electro-optical boundary in today's global telecommunications networks. These technologies, however, are perhaps better suited for long-haul applications and possibly for some metro core applications. The possibility of cost-effective use in metro access remains to be demonstrated.

  The ODSI specification is the industry's first electrical-optical interface protocol to support IP routing over optical layers of the public network. The coalition has created and written code for a specification that brings cooperation among access, edge, and core devices, and also supports signaling for legacy third-party devices. The goal of the ODSI coalition is to hand off its work to the industry's established standard-setting bodies such as the Optical Internetworking Forum (OIF), the Internet Engineering Task Force (IETF), the ITU, and T1/X1. In late 2000, the ODSI coalition announced the successful completion of the first multivendor interoperability testing of the ODSI specification. Alcatel, Equipe Communications, Redback Networks, Sycamore Networks, and Tenor Networks demonstrated a successful multivendor implementation of ODSI's standards-based protocol specification between various IP and optical devices.

  Optical User-Network Interface (O-UNI) and Optical Network-Node Interface (O-NNI) were introduced in Chapter 2 and are covered in more detail in Chapter 8.

End Notes

1. You can also carry a non-SONET framed signal; however, the majority of systems now deployed are used as adjuncts to SONET—namely, they carry a SONET signal.
2. It is also possible to pack both a send and a receive signal onto a single strand, but most telecom systems utilize a fiber pair.

3. The methodologies embodied in these trade-off considerations are really no different from the trade-off considerations of the 1950s when contemplating the deployment of a 12-channel analog loop carrier as contrasted with the deployment of 12 distinct analog loops, or from considerations in the 1970s when contemplating the deployment of a 24-channel digital loop carrier as contrasted with the deployment of 24 distinct (analog) loops.

4. For example, note this quote from J. McQuillan, "Broadband Networking at the Crossroad," NGN 2001 Promotional Materials, October 11, 2001: "Ten years of price competition in voice have been compressed into three years for Internet prices. And price erosion is accelerating: A contract for an OC-3 from New York to Los Angeles for the year 2002 dropped in price by 92 percent in just four months."

5. T1 and T2 systems have been more predominant in the loop environment than T3 systems; however, with the introduction of fiber-based feeder systems, particularly the Bell System/Lucent SLC-96 system, an increasing number of DS3 systems (originally called FT3) have been deployed in support of loop applications.

6. Refer to Bellcore, *Telecommunications Transmission Engineering*, ISBN 1-878108-04-2 (1990): 396 ff.

7. www.cir-inc.com/index.cfm?loc=news/view&id=85. Excerpting from the November 2001 report by Communications Industry Researchers, Inc. (CIR) titled "Metro Optical Networking Market Opportunities," which states

> Are we there yet? For close to five years, the metro DWDM market has failed to fulfill the vaunted promises of significant growth opportunities. CIR's research prognosticates that 2003 will be the first big year for substantial metro DWDM deployment to occur in the United States, with the market reaching $500 million in 2005. CIR predicts Verizon will be the biggest buyer of metro DWDM ring equipment among the RBOCs, with the regulated portion of Qwest being the next largest user, followed by more staunch SONET multiprotocol gear supporters, SBC and BellSouth. Simply put, metro optical networking market development will remain stuck in neutral until we see more widespread acceptance of DWDM ring technology by the RBOCs and other metro carriers. It has been the classic chicken and the egg scenario since early 1997, in which the industry has been waiting for a critical mass of wavelengths in the metro space to enable both the OADM and optical switching markets to develop. But, until the vendors can provide a combination of lower cost and engineering-friendly DWDM products with true ring capability, the

all-optical vision in the metro segment will not begin to be realized. CIR believes that the truth about the metro space reflects the economic realities of today, and not the market hype indicating that there are quick-fix elixirs. Most carriers will continue to consolidate spending and look for operational and cost efficiencies. The mergers between the different RBOCs and GTE over the past five years have left large amounts of room for carriers to reduce head counts as well as to consolidate networks and organizations. Many of the emerging carriers that burst onto the scene with great fanfare promoting better, faster, and cheaper services are now on shaky financial footing. The overall economy's slowdown will translate into slowing demand for services, which means that incumbent carriers will strive to get by with as much of their legacy gear as they can as they work through the correction. CIR believes that companies such as Lucent, Nortel, Fujitsu and Cisco will most likely dominate the space based on their installed base and sheer size. Riverstone will lead Ethernet vendors, but by balancing Ethernet startup providers with international accounts and OEM arrangements. Cisco Capital is ensuring that Cisco remains on top with CLECs by continuing to finance startups like Cogent and Sigma. ONI could possibly be a primary U.S. metro core player at some point, pending a successful completion of the OSMINE process from Telcordia, but Tellabs and Ciena are likely to struggle in this space for the foreseeable future. The RBOCs will probably continue to dominate the market for metro core optical services, instead of cash-strapped and debt-laden service providers, such as XO, MFN, Level 3, Williams and others. Therefore, CIR predicts a much more moderate growth rate of optical gear purchases in the space, based on the historical rate of service introduction by companies like Verizon, SBC, and BellSouth.

8. Phased-array WDMs are imaging devices. They image the field of an input waveguide onto an array of output waveguides in a dispersive way. In phased-array-based devices, the imaging is provided by an array of waveguides, the length of which has been chosen to obtain the required optical/dispersive properties. Phased-array-based devices are simpler and more robust than other devices (such as grating-based devices, which require vertical etching to obtain a vertically etched reflection grating).

9. For example, see S. I. Najafi et al., "Passive and Active Sol-Gel Materials and Devices," *SPIE Press* CR68 (1997): 253–285; S. I. Najafi et al., "Ultraviolet Light Imprinted Sol-Gel Silical Glass Waveguide Devices," *Optics Communications* (1996): 128; S. I. Najafi et al., "Sol-Gel Integrated Optics Coupler by Ultraviolet Light Imprinting," *Electronics Letters* (1995): 31; S. I. Najafi et al., "Ultraviolet Light Imprinted Sol-

Gel Silical Glass Low-Loss Waveguides for Use at 1,550 Nanometers," *Optical Engineering* (1997): 36; S. I. Najafi et al., "Gratings Fabrication by Ultraviolet Light Imprinting and Embossing in Sol-Gel Silica Glasses," *Proceedings of SPIE* 2695 (1996); M. P. Andrews, et al., "Erbium-Doped Sol-Gel Glasses for Integrated Optics," *Proceedings of SPIE* 2397 (1995); S. I. Najafi et al., "Erbium in Photosensitive Hybrid Organoaluminosilicate Sol-Gel Glasses," *Proceedings of SPIE* 2997 (1997); S. I. Najafi et al., "Fabrication of Ridge Waveguides: A New Sol-Gel Route," *Applied Optics* (1997); S. I. Najafi et al., "UV-light Imprinted Bragg Grating in Sol-Gel Ridge Waveguide with almost 100% Reflectivity," *Electronics Letters* (1997); S. I. Najafi et al., "Photoinduced Structural Relaxation and Densification in Sol-Gel-Derived Nanocomposite Thin Films: Implications for Integrated Optics Device Fabrication," *Canandian Journal of Chemistry* (November 1998).

10. However, opportunities also exist for combining these techniques with monolithic, semiconductor-based, optoelectronic, multichip module techniques to bring forth complex, high-performance, and low-cost data and telecommunications modules.

11. Cisco Sources and CIR Inc. et al., NGN Ventures 2001.

12. R. A. Barry, "Intelligent Optical Networking Year 2001: A Practical Perspective," NGN 2001 Proceedings, 2001.

13. Brent Allen and Solomon Wong, "Is WDM Ready for Local Networks?" Nortel White Paper, Document 56005.25-05-99, 1997.

14. For the purpose of sizing the market in the previous discussion, however, we assumed that the entire system was deployed on day 1.

15. Transmission systems at 40 Gbps are now the next frontier in optical communication. At the transmitter end, appropriate source modulation is critical to deal with effects such as fiber dispersion and nonlinearity. Modulation requirements for efficient transmission at 40 Gbps impacts chirp, drive voltages, insertion loss, extinct ratio, power-handling capability, and wavelength dependence. Components such as LiNbO3, GaAs Polymer, and integrated electroabsorption modulators may play a role in this context.

16. Earlier in the chapter, we discussed a 60-city universe; the model of Chapter 1 considers a 90-city universe.

17. The carrier for whom we perform strategic planning assumes that the carrier will penetrate the top 7,000 buildings (100 in year 1, 900 more in year 2, and so on) and will achieve an average of 3 broadband customers per building. This conservative financial approach to the matter fits well with the result of this table.

18. This is based on the order flow seen at a MASP.
19. As discussed in a number of places in this book, cost-effective NEs are one component of a viable broadband carrier undertaking. Other components include reduced operational and fiber spur costs.
20. This is similar to the example discussed earlier with the sol-gel approach.

CHAPTER 6

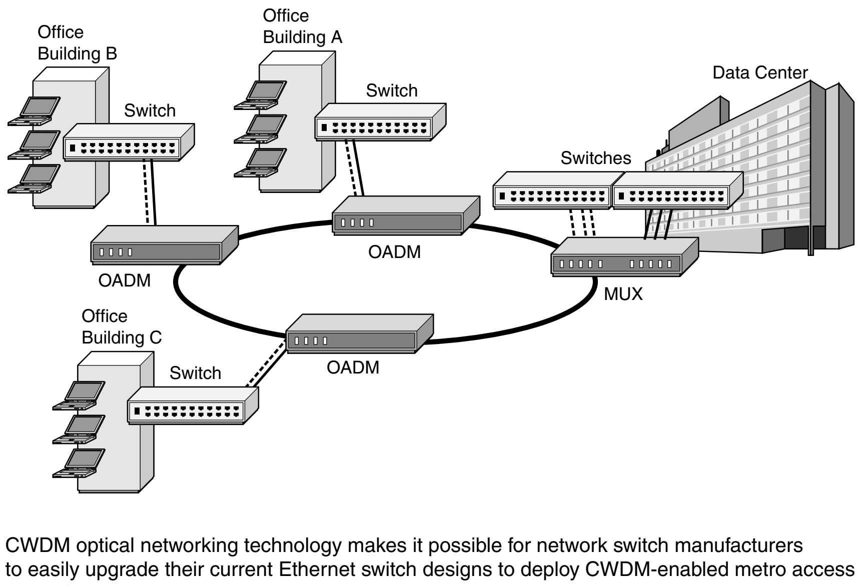
Coarse Wavelength Division Multiplexing (CWDM) Technology

This chapter covers coarse wavelength division multiplexing (CWDM), which appears to be a reasonably good technology for metro access applications if certain extensions are developed by vendors; however, it has not yet experienced widespread use. We also cover some related technologies in this chapter.

6.1 Technology Scope

In principle, CWDM can be used by metropolitan area carriers that want to keep the cost per building low, so that they can turn a positive net bottom line. Figure 6-1 shows a simple application of the technology, although more sophisticated systems and designs are possible.[1] CWDM technology may be used effectively until either the cost of dense WDM (DWDM) remains higher than that of CWDM or the bandwidth requirement per enterprise building exceeds 10, 40, or 320 Gbps. We discussed in Chapter 1, "Advances and Opportunities in Next-Generation SONET and Other Optical Architectures," that a typical 16,000-square-foot office building in the United States currently needs about 10 Mbps and will need about 73 Mbps by

Figure 6-1 Example of a simple application of CWDM.



CWDM optical networking technology makes it possible for network switch manufacturers to easily upgrade their current Ethernet switch designs to deploy CWDM-enabled metro access networks.

2006. Therefore, it appears that unless the cost of DWDM decreases significantly, CWDM has a window of opportunity for the immediate future in the metropolitan area network (MAN) space. It is going to be a challenge to deploy a bottom-line positive metro access/metro core operation. Among other things, the cost per building has to be kept low. Optical equipment vendors that have developed long-haul DWDM equipment will find that they cannot simply relabel the product by taking a long-haul DWDM system and repainting the word “metro” over “long-haul” to secure a cost-effective metro DWDM product. However, some advancement will be needed before CWDM technology becomes poised for wide-scale deployment. In particular, it is crucial that CWDM systems be enhanced with strong network management/operations support system (OSS) tools, so that networks based on these network elements can be supported cost-effectively in terms of provisioning, alarming, and fault restoration.

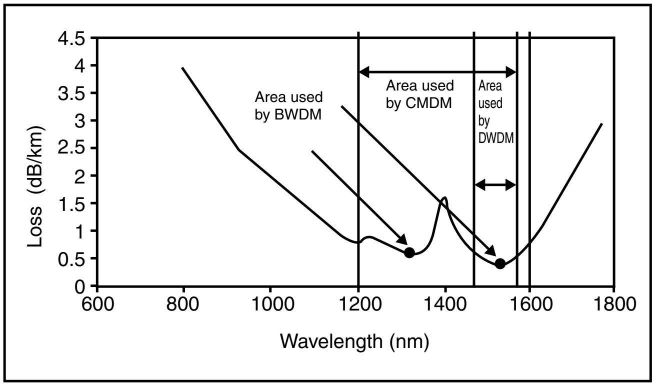
CWDM is an implementation of WDM technology that is applicable to metro access applications and short- to medium-haul networks. Because CWDM uses low-cost uncooled lasers and wideband multiplexers and demultiplexers instead of the more expensive cooled lasers found in standard DWDM systems, it provides a much more affordable solution.

Typically, there is no immediate need for 32 channels in the access network. From a power budget standpoint, designers may want to limit themselves to 8 or 16 buildings per ring. Utilizing 32 lambdas would equate to 4 lambdas per building. Providing 40 Gbps (4×10 Gbps) or 10 Gbps (4×2.5 Gbps) is currently overkill. This simply builds an expensive and nonrevenue-producing inventory of capacity that will not be able to be utilized in a money-making manner for the foreseeable future. Therefore, it makes sense to provide an inexpensive way for carriers to use WDM. Low-cost transceiver optics along with inexpensive passive components for CWDM using 8, 12, or 16 channels make this approach a viable solution. CWDM technology can be deployed to take advantage of this evolving metro opportunity and deliver significant cost savings to local area networks (LANs), storage area networks (SANs), and metro access broadband optical carrier networks. End users, in turn, can access, accumulate, and disseminate information over these networks more cost-effectively. As noted, however, one of the critical issues yet to be fully addressed is Operations, Administration, Maintenance, and Provisioning (OAM&P) support. Unfortunately, CWDM technology has not yet developed to the point where a well-recognized architecture for standard deployment is available and a set of operational disciplines in the OAM&P context (supported by feature-rich OSSs) exists. Hence, the de facto alternative to CWDM-based metropolitan systems at this time is next-generation SONET.

One way that CWDM and DWDM differ noticeably is in the spacing between adjacent wavelengths. DWDM packs many channels into a small usable spectrum, spacing them one to two nanometers (nm) apart. DWDM systems support a high channel count, but also require expensive cooling equipment and independent lasers and modulators to ensure that adjacent channels do not interfere. CWDM systems, on the other hand, use 10- to 25-nm spacings with 1,310- or 850-nm lasers that drift less than 0.1 nm/° centigrade (C). This low drift eliminates the need for cooling equipment, which, in turn, reduces the total system cost. As a result, CWDM systems support less total bandwidth than DWDM systems; however, with 8 to 16 channels, each operating between 155 Mbps and 2.5 Gbps (and soon 10 Gbps), these systems can deliver bandwidths ranging from 1 Gbps to over 100 Gbps. Figure 6-2 plots the general wavelengths used by various WDM technologies.[2] Typical systems support 8 wavelengths, data rates up to 2.5 Gbps per wavelength, and distances up to 50 km.

CWDM uses lasers with a wide channel wavelength spacing. In contrast, DWDM, which is widely used in long-haul networks and some metro core

Figure 6-2
Wavelengths used by WDM.



- DWDM refers to the ability to carry a large number of channels, where large may be a hundred or more. DWDM equipment usually uses the 1530- to 1565-nm wavelength range with very high laser accuracy requiring temperature-controlled components. The standards today define a 0.8-nm wavelength gap between any two channels, with some manufacturers demonstrating equipment with 0.4-nm and even 0.2-nm accuracy.

- CWDM uses a much wider wavelength range (1200 to 1600 nm) with a minimum 20-nm wavelength gap between any two channels. Optical components used for CWDM are less accurate and thus less expensive because of the 20-nm wavelength gap.

- Bidirectional WDM (BWDM) refers to the ability to carry two channels in different directions on a single fiber. Two distinct wavelengths may be used in BWDM technology: 1310 and 1550 nm.

networks (particularly those with large diameters), uses lasers with much narrower wavelength spacing, typically 0.8 or 0.4 nm. The wide channel spacing of CWDM means that a lower system cost can be achieved. This lower equipment cost is the result of a lower transmitter cost because no temperature and wavelength control is needed and a lower optical mux/demux cost is due to wider tolerance on the wavelength stability and bandwidth.

Transmission-side technological advancements are currently taking place. For example, Lucent recently demonstrated the highest-capacity (20-Gbps) transmission on an access fiber (AllWave™ zero water peak) using O and E bands (1,300- to 1,440-nm band), and 8-channel, 20-nm-spaced CWDM technology; each channel supports 2.5 Gbps. This kind of full-spectrum-capable CWDM solution provides the lowest cost per bit and is upgradeable to DWDM in the future. This CWDM technology is based on a 20-nm channel spacing that allows lower-cost components and systems over a range from 1,490 to 1,610 nm.

It is important to emphasize the unique requirements of metropolitan applications, particularly the add/drop ring aspect of metro local traffic compared to more linear applications in long haul. Some vendors are combining not only appropriate edge network element (NE) technology, but also waveguide technology—for example, Lucent TrueWave RS fiber. TrueWave RS works with off-the-shelf commercial lasers ensuring high laser yield, system interoperability, and multiple vendor sourcing—issues that ensure low-cost systems for service providers. Proponents make the case that CWDM solutions on AllWave are expected to yield an equipment cost savings of 40 percent compared to DWDM over single-mode fiber. These savings are achieved by using uncooled lasers and relaxing stringent DWDM filter specifications, and will result in an improvement in service provider gross profitability. Table 6-1 depicts key parameters for a typical CWDM system. The bottom line is that adding two to eight wavelengths per fiber in metro short-haul applications enables network designers to increase capacity without installing costly DWDM systems.

CWDM represents significant costs savings—from 25 to 50 percent at the component level over DWDM, for original equipment manufacturers (OEMs) and service providers. At the time of this writing, CWDM products cost about $3,500 per wavelength. That is half the price of regular DWDM products, which typically sell for $7,500 per wavelength. Traditional CWDM only scales to about eight wavelengths, but for metro access applications, this may be adequate. Also, vendors have found ways to combine CWDM with regular DWDM blades that enable the system to scale up to 20 wavelengths. CWDM system architecture can benefit the

Table 6-1

Key Parameters for a Typical CWDM System

CWDM Series		Unit
Center wavelength λc	1,471.0, 1,491.0, 1,511.0, 1,531.0, 1,551.0, 1,571.0, 1,591.0, 1,611.0	nm
Passband bandwidth	λc +/− 6.5	nm
Insertion loss	<4.5	dB
In-band ripple	<0.5	dB
Polarization dependent loss (PDL)	<0.1	dB
Adjacent band isolation	>25	dB
Return loss without connectors	>45	dB
Directivity	>55	dB
Operating power	300	mW
Operating temperature	0–70	°C
Storage temperature	−40–85	°C
Fiber type	Corning SMF-28	
Fiber package	900μm loose tube	

metro access market because it takes advantage of the inherent natural properties of the optical devices and eliminates the need to artificially control the component characteristics.

The following paragraphs identify typical CWDM optical elements:

CWDM uncooled coaxial lasers. Distributed-feedback multiquantum well (DFB/MQW) lasers are often used in CWDM systems. These lasers typically come in 8 wavelengths and feature a 13-nm bandwidth. Wavelength drift is typically only 5 nm under normal office conditions (say, with a 50°C total temperature delta), making temperature compensation unnecessary. For additional cost savings, the lasers do not require external gratings or other filters to achieve CDWM operation. They are available with or without an integral isolator.

CWDM transmitters/receivers. OC-48 CWDM transmitters typically use an uncooled DFB laser diode (LD) and are pigtailed devices in a standard 24-pin dual inline package (DIP). Six to 8 channels are supported

(6 channels are located at 1,510 to 1,610 nm, and 2 additional channels are located at 1,470 and at 1,490 nm). The OC-48 receiver typically uses an avalanche photodetector diode (APD) photodetector, has a built-in DC-DC converter, and employs a phase-locked loop (PLL) for clock recovery. Transmission distances of up to 50 km are achievable with these modules.

CWDM wavelength division multiplexers/demultiplexers. In four- or eight-channel modules, you would typically use thin-film filters optimized for CWDM applications, with filtering bands matched to the CWDM wavelengths. Filters need to feature low insertion loss and high isolation between adjacent channels.

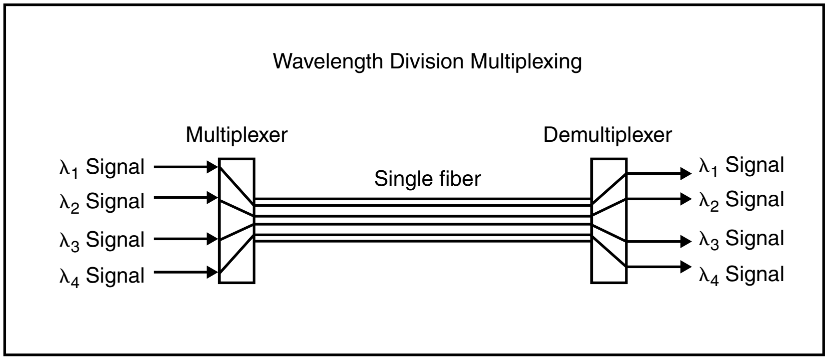
CWDM optical add/drop modules (OADMs). These are available in various configurations with one, two, or four add/drop channels using the same thin-film filters as the CWDM mux and demux modules.

6.2 CWDM Systems

CWDM[3] technology was first commercially deployed in the early 1980s for transporting digital video signals over multimode fiber. Quante Corporation marketed a 4-wavelength system that operated in the 800-nm window with 4 channels, each operating at 140 Mbps. These systems were primarily used in cable TV links. However, CWDM systems did not generate significant interest among service providers—that is, until now. As metro carriers seek cost-effective solutions to their transport needs, CWDM is becoming more widely accepted as an important transport technology. As noted, unlike DWDM, systems based on CWDM technology utilize uncooled DFB lasers and wideband optical filters. These optical elements provide several advantages to CWDM systems such as lower power dissipation, smaller size, and lower cost. The commercial availability of CWDM systems offering these benefits makes the technology a viable alternative to DWDM systems for many metro and access applications.

The bandwidth of a fiber-optic link can be increased by transmitting data faster or transmitting multiple wavelengths in a single fiber. As discussed at length in Chapter 5, "DWDM Systems," WDM is accomplished by using a multiplexer to combine wavelengths traveling on different SONET-managed fibers into a single fiber. At the receiver end of the link, a demultiplexer separates the wavelengths and routes them into different fibers, which all terminate at separate receivers (see Figure 6-3). The spacing between the individual wavelengths transmitted through the

Figure 6-3
WDM revisited.



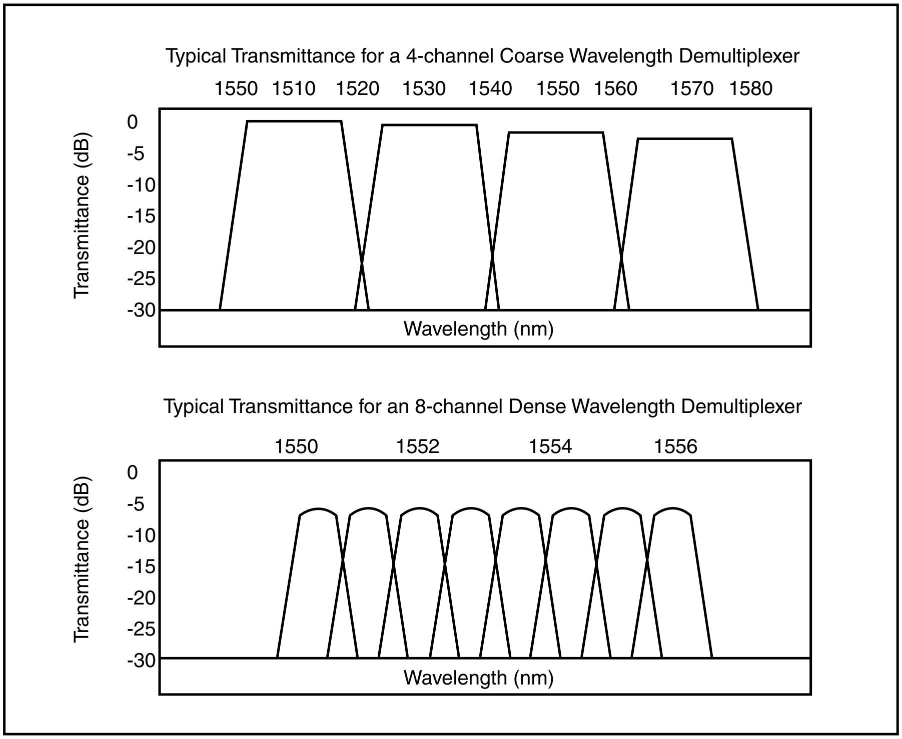
WDM is accomplished using a multiplexer to combine wavelengths traveling on separate fibers into a single fiber. At the receiver end of the link, a demultiplexer separates the wavelengths and routes them into different fibers.

same fiber serves as the basis for differentiating DWDM from CWDM (see Figure 6-4).

DWDM systems typically use wavelength separations of 200 GHz (1.6 nm), 100 GHz (0.8 nm), or 50 GHz (0.4 nm), with future systems projected to have even narrower spacing at 25 GHz or less. The operating wavelengths of DWDM systems are defined according to a standardized frequency/wavelength grid developed by the International Telecommunications Union (ITU), which was discussed in Chapter 5.

DFB lasers are used as sources in DWDM systems. The laser wavelength drifts about 0.08 nm/°C with the temperature. DFB lasers are cooled to stabilize the wavelength from drifting outside the passband of the multiplexer and demultiplexer filters as the temperature fluctuates in DWDM systems. CWDM systems use DFB lasers that are not cooled. They are specified to operate from 0 to 70°C with the laser wavelength drifting only about 5 to 6 nm over this range (for example, 0.1 nm/°C). This wavelength drift coupled with the variation in laser wavelength of up to ±3 nm due to laser die manufacturing processes yields a total wavelength variation of about 12 nm. The optical filter passband and laser channel spacing must be wide enough to accommodate the wavelength variation of the uncooled lasers in CWDM systems (see Figure 6-5). Channel spacing in these systems is typically 20 nm with a channel bandwidth of 13 nm. CWDM systems offer some key advantages over DWDM systems for applications that require channel counts on the order of 16 or less. These benefits include better costs, power requirements, and size.

Figure 6-4
Lambda spacing.

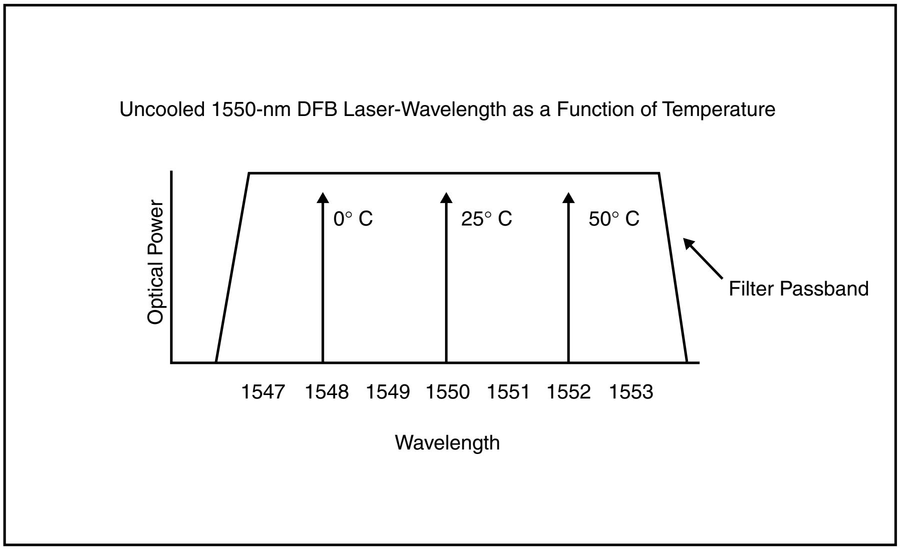


The spacing between the individual wavelengths transmitted through the same fiber serves as the basis for differentiating dense wavelength division multiplexing from coarse WDM (CWDM). Spacing in CWDM systems is typically 20 nm, while most DWDM systems today offer 0.8 nm (100 GHz) wavelength separation.

6.2.1 Less Expensive Hardware

The cost difference between CWDM and DWDM systems can be attributed to hardware and operating costs. Although DWDM lasers are more expensive than CWDM lasers, the cooled DFB lasers provide cost-effective solutions for long-haul transport and large metro rings requiring high capacity. In both of these applications, the cost of the DWDM systems is amortized over the large number of customers served by these systems. Metro access networks, on the other hand, require lower-cost and lower-capacity systems to meet market requirements, which are based largely on what the customer is willing to pay for broadband services.

Figure 6-5
CWDM drift.



Uncooled coarse WDM lasers drift in wavelength at the rate of approximately 0.08 nm/centigrade. The filter passband and laser channel spacing must be wide enough to accommodate this temperature dependent wavelength drift, plus any manufacturing variations.

Primary market research shows that customers typically want to spend at most $4,000, $7,000, or $12,000 per month for a dedicated 10 Mbps, 100 Mbps, or 1 Gbps service for a two-point metropolitan transparent LAN service (TLS), respectively. The backbone fiber and the spurs (laterals) in the two buildings cost about $4,000 per month. The carrier's operations and sales, general, and administrative (SGA) costs are in the 40 percent range, or $1,600, $2,800, and $4,800 per month for the aforementioned 10 Mbps, 100 Mbps, and 1 Gbps services, respectively. This leaves a monthly charge of $1,600 for 10 Mbps, $200 for 100 Mbps, and $3,200 for 1-Gbps edge equipment at two building sites and central office (CO) equipment. Ignoring for a moment the CO equipment and assuming a 25-month payback on the edge equipment, the carrier can only invest $0[4] at each site for 10-Mbps TLS, $5,000 for 100-Mbps TLS ($2,500 per end), and $80,000 for 1-Gbps TLS ($40,000 per end). In addition, you need to add in some profit for the carrier and the carrier's shareholders. Currently, there are few customers who need 1 Gbps in the WAN; therefore, the sweet spot for the carrier is 100-Mbps TLS. As a result, the equipment cost has to be in the range of $2,500 per building. These figures clearly throw the design in the direction of CWDM (or CWDM-like price points).

A metro carrier needs to find more than one customer in a building to be viable. Although some carriers have publicly stated that they need/want 67 broadband (FE/GbE) customers per building, we find this totally untenable. A much more reasonable assumption is that a carrier will find three customers per building. This would open up the budget for the edge equipment to some extent, but when you also consider the CO costs and the required profit figure (which needs to be at least 15 to 25 percent of the gross), you are left, once again, with some tight requirements on the building equipment, throwing the design in the direction of CWDM-like price points.

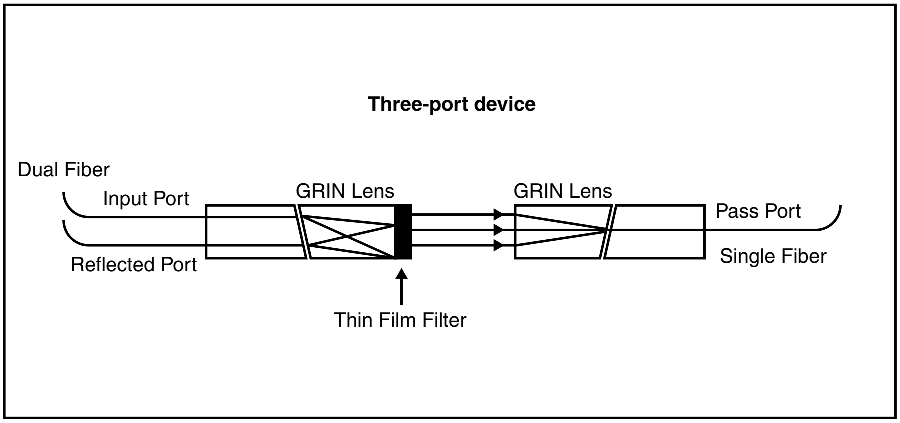
The price of DWDM transceivers is typically four to five times more expensive than that of their CWDM counterparts. The higher DWDM transceiver costs are attributed to a number of factors related to the lasers. The manufacturing wavelength tolerance of a DWDM laser die compared to a CWDM die is a key factor. Typical wavelength tolerances for DWDM lasers are on the order of ±0.1 nm, whereas manufacturing wavelength tolerances for a CWDM laser die are ±2 to 3 nm. Lower die yields thus drive up the costs of DWDM lasers relative to CWDM lasers. In addition, packaging a DWDM laser die for temperature stabilization with a Peltier cooler and thermister in a butterfly package is more expensive than uncooled CWDM coaxial laser packaging.

The cost difference between DWDM and the CWDM multiplexers and demultiplexers based on thin-film filter technology contributes to lower overall system costs in favor of CWDM as well. CWDM filters are inherently less expensive to manufacture than DWDM filters due to the fewer number of layers in the filter design. Typically, there are approximately 150 layers for a 100-GHz filter design used in DWDM systems, whereas there are approximately 50 layers in a 20-nm CWDM filter. The result is higher manufacturing yields for CWDM filters (see Figure 6-6). The CWDM filter costs about 50 percent less than the DWDM filter, and the cost is projected to drop by a factor of three in the next two to three years as automated manufacturing takes hold. The adoption of new filter and multiplexing/demultiplexing technologies is expected to decrease costs even further.

6.2.2 Lower-Power Requirements

The operating costs of optical transport systems depend on maintenance and power. Although maintenance costs are assumed to be comparable for both CWDM and DWDM systems, the power requirements for DWDM are significantly higher. For example, DWDM lasers are temperature-stabilized with Peltier coolers integrated into their module package. The cooler along

Figure 6-6
CWDM filter designs.



Coarse WDM (CWDM) filter designs use fewer layers, and the alignment tolerances are relaxed relative to those for DWDM devices, which significantly lowers the manufacturing costs of the CWDM devices.

with the associated monitor and control circuitry consumes around 4 watts per wavelength. Meanwhile, an uncooled CWDM laser transmitter uses about 0.5 watts of power. The transmitters in an 8-channel CWDM system consume approximately 4 watts of power, while the same functionality in a DWDM system can consume up to 30 watts. As the number of wavelengths in DWDM systems and transmission speeds increases, power and the thermal management associated with them become a critical issue for board designers. The lower-power requirement resulting from the use of uncooled lasers in CWDM systems has positive financial implications for system operators. For example, the cost of battery backup is a major consideration in the operation of transport equipment. Minimizing operating power and the costs associated with its backup, whether in a CO or a wiring closet, reduces operating costs.

6.2.3 Comparable Reliability

The reliability of DFB lasers used in DWDM and CWDM transport has been proved in both cooled and uncooled designs. The difference between the two laser designs is the number of additional components, including the Peltier cooler, thermister, and associated electronics in DWDM lasers. However, manufacturers generally claim that no reliability difference exists between the two types of systems in real-world applications.

6.2.4 Smaller Physical Size

CWDM lasers are significantly smaller than DWDM lasers. Uncooled lasers are typically constructed with a laser die and monitor photodiode mounted in a hermetically sealed container with a glass window. These containers are aligned with a fiber pigtail or an alignment sleeve that accepts a connector. The container and sleeve form a cylindrical package called a Transmitter Optical Subassembly (TOSA). A typical TOSA is approximately 2 cm in length and 0.5 cm in diameter. Cooled lasers are offered in either a butterfly or dual-inline laser package and contain the laser die, monitor photodiode, thermister, and Peltier cooler. These lasers are about 4 cm long, 2 cm high, and 2 cm wide. These devices are almost always pigtailed, requiring fiber management, a heatsink, and a corresponding monitor and control circuitry. The size of a DWDM laser transmitter typically occupies about five times the volume of a CWDM transmitter.

6.2.5 Up to 16 Wavelengths

CWDM systems supporting two to eight wavelengths have been commercially available for some time. Newer systems (see Figure 6-7) now scale to

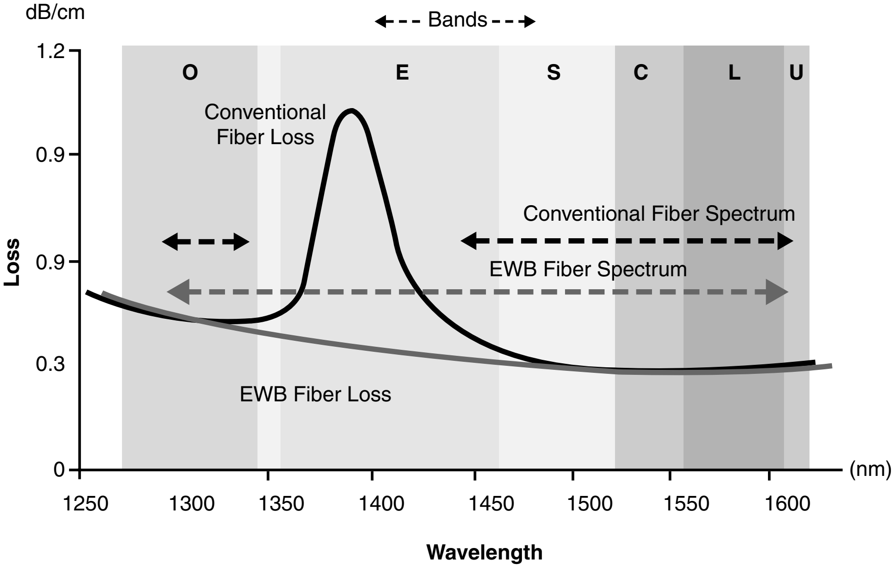


Figure 6-7
Optical bands.

16 wavelengths in the 1,290- to 1,610-nm spectrum. Today most CWDM systems are based on 20-nm channel spacing from 1,470 to 1,610 nm with some development occurring in the 1,300-nm window for 10 Gigabit Ethernet (GbE). With conventional fiber, wavelengths in the 1,400-nm region suffer higher optical loss due to the attenuation peak caused by residual water present in most of the currently installed fiber. Although this additional loss can limit system performance for longer links, it is not an obstacle to CWDM deployment in most metro access spans. New fiber that eliminates the water attenuation peak is offered by at least two of the primary fiber vendors for use in metro links with loss budgets that mandate lower fiber attenuation. Extended wavelength band (EWB) fiber easily supports 16 channels, whereas conventional single-mode fiber (SMF) easily supports 12 channels.

The advantages of EWB fibers for CWDM applications are as follows:

EWB Fiber

EWB fiber has identical specifications to standard single-mode fiber plus the following:

- 25 percent fewer fibers in the CWDM feeder
- 25 percent less CWDM CO floor space and apparatus requirements
- Improved CWDM packaging economy
- Lowers installed CWDM system cost by 10 percent
- Opens E-band spectrum
- Maximum future proofing for a 25-year or more fiber plant

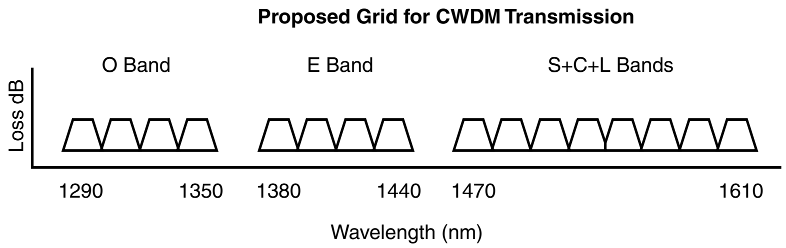
Long-wavelength InGaAs/InP avalanche photodiodes are helping make low-cost systems a reality. The trade-off in many low-cost systems is the link budget. Although many transceivers are capable of coming close to meeting the performance specifications of long-haul equipment, the passive optical losses induced by couplers, splitters, multiplexers, and demultiplexers often overburden the optical link budget. In order to keep the system low cost, expensive erbium-doped fiber amplifiers (EDFAs) are not utilized;

instead, you can make use of long-wavelength avalanche photodiode technology. InGaAs/InP avalanche photodiodes have been used to increase the receiver sensitivity of long-haul communication systems. They provide current gain on the order of 10 times in the receiver front end, prior to the transimpedance amplifier, which boosts the signal before being received. In the S, C, and L bands, it is necessary to use a separate absorption and multiplication layer structure, wherein the light is absorbed in a layer of InGaAs and the resulting signal is amplified in an InP multiplication layer. Unfortunately, avalanche photodiodes have been plagued historically by low-yield processes and a reputation of not being as reliable as standard PIN diodes. Additionally, the need to provide a high-voltage supply (40 to 60 volts) and track the APD supply voltage over temperature complicated their use in many systems. Improvements in epitaxial growth and processing techniques have enabled the vendors to achieve high yields (90 percent), with reliabilities equivalent to PIN diodes.[1]

6.2.6 Standards Underway

CWDM systems development and standards efforts come at a critical time for metro access service providers. Work is underway for the transport plane, but additional work is needed in the management plane, as noted earlier. One organization working to define standards for CWDM systems in the 1,400-nm region is the Commercial Interest Group (CIG). CIG participants include component suppliers, system vendors, and service providers. The group's focus to date has been primarily on defining the CWDM wavelength grid and investigating the cost/performance comparisons of DWDM versus CWDM architectures. The CWDM wavelength grid under consideration for proposal is divided into three bands. The O band consists of 1,290, 1,310, 1,330, and 1,350 nm. The E band consists of 1,380, 1,400, 1,420, and 1,440 nm. The S, C, and L bands consist of 8 wavelengths from 1,470 to 1,610 nm in 20-nm increments. These wavelengths take advantage of the full optical fiber spectrum, including the legacy optical sources at 1,310, 1,510, and 1,550 nm, while maximizing the number of channels. The 20-nm channel spacing supports lower component costs with the use of uncooled lasers and a wideband filter. It also avoids the high-loss 1,270-nm wavelength and maintains a 30-nm gap for adjacent band isolation (see Figure 6-8). Other standardization activities are underway, as shown in the following text.[5]

Figure 6-8
Standards proposal.



The Commercial Interest Group is working to define a standard CWDM wavelength grid.

CWDM Grids Proposed to the Telecommunications Industry Association (TIA) (1/01) and ITU (2/01)

Coarse channel plan 1:

O band	1,270, 1,290, 1,310, and 1,330 nm
E band	1,370, 1,390, 1,410, and 1,430 nm
S, C, and L bands	1,470, 1,490, 1,510, 1,530, 1,550, 1,570, 1,590, and 1,610 nm

Coarse channel plan 2:

O band	1,275.7, 1,300.2, 1,324.7, and 1,349.2 nm (802.3ae 10 GbE)
E band	1,380, 1,400, 1,420, and 1,440 nm
S, C, and L bands	1,470, 1,490, 1,510, 1,530, 1,550, 1,570, 1,590, and 1,610 nm

CWDM Grid Standards Activities

- The Institute of Electrical and Electronics Engineers (IEEE) (P802.3ae) has proposals to support four CWDM channels in the O band (1,260 to 1,360 nm)
- January 2001 approvals
  - TIA approval of proposed grids as a Lucent contribution to ITU
  - U.S. Department of Commerce SGB approval as a Lucent contribution to ITU
- February 2001 presentation at ITU
  - **Q.16 systems** Directed to present "white contribution" at next conference in October, and proposed as a new recommendation or as a change to a previous recommendation
  - **Q.17 components** Identified new CWDM classification for WDM components for G.671 update

As the demand for bandwidth is pushed to the edge of the network, the need for low-cost transport systems is imperative. CWDM technology fits these requirements, offering a scalable system architecture for metro and access networks.

6.3 Other Applications

Three other applications are identified in this section for CWDM technology:

- LAN applications
- First-mile applications
- Poor man's metro networks

For some LAN applications, CWDM is an attractive option for increasing the data throughput while utilizing the installed multimode fiber (MMF) base. CWDM modules based on 840-nm wavelength Vertical Cavity Surface-Emitting Lasers (VCSELs) or 1,310-nm DFB lasers have already been developed, enabling 4×2.5 Gbps transmission over 100-m MMF or 300-m

MMF, respectively, both with a 62.5-μm core diameter. However, recent demonstrations of 10-Gbps VCSEL transmitters and MMF receivers together with the progress of silicon germanium (SiGe) electronics show the feasibility of a straightforward, serial, high-speed solution without the added complexity of optical multiplexing and demultiplexing. Indeed, 10-Gbps data transmissions over record distances of 1.6 km or 2.8 km of a new high-bandwidth 50-μm core diameter MMF have been reported recently.[6,7] Subcarrier multiplexing and multilevel coding as alternative upgrading methods have not yet been demonstrated in sufficient quality and are not considered currently competitive.[8] Likewise, today's parallel optical modules, already with an aggregate data rate up to 30 Gbps through 12 channels, seem unsuitable due to the high cost of a few 100-m-long fiber ribbon cable.[9] With a further increase of bandwidth demand beyond the direct current modulation capabilities of laser transmitters, CWDM approaches will gain importance for building backbone links. Researchers have already demonstrated a 4-channel, 10-Gbps CWDM link over 310 m of high-performance MMF cable.[10,11]

The second application is in support of Ethernet in the First Mile (EFM) work of the IEEE. Figure 6-9 shows a CWDM application in the point-to-point Ethernet mux application,[5] and Figure 6-10 shows a CWDM application in the Ethernet passive optical network (EPON).[5]

Others advocate what we call a *poor man's metropolitan network*, which is a network achieved by extending the optics of GbE to cover longer distances on the assumption that dark fiber can be secured by an end user who will assume the economic burden. Typically, these networks are not carrier class, do not have the requisite security, and are difficult to scale and manage.

This approach uses extended-optic gigabit interface converters (GBICs). Although GBIC slots in today's switches were designed for different reasons, they still have the potential to be expanded into access points for low-grade MANs. In this scenario, according to proponents, you should deploy fiber rings within cities and connect them to Ethernet switches via their GBIC slots to create high-bandwidth MANs at a low cost using CWDM transceivers and optical multiplexers.[1]

In this design, an optical fiber ring is installed in a building with fibers from both the east and west linking with other buildings. Data is transmitted at various wavelengths on the ring and can be pulled off or added to the ring at various points or nodes to remove and insert data into the network through an optical add/drop multiplexer (OADM). Data can travel both east and west around the ring to provide two redundant pathways, so that the network remains viable even if the fiber is broken at one point on the ring. Nodes along the ring can communicate directly with each other via virtual

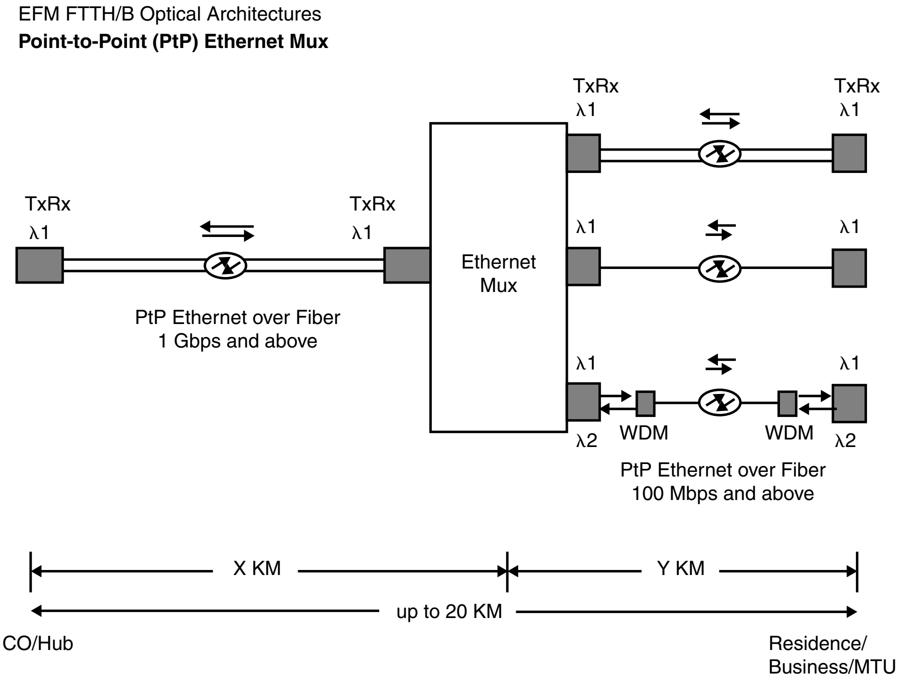


Figure 6-9 EFM FTTH/B optical architecture.

point-to-point connections that pass through the intermediary devices. Even though data is flowing around a physical ring, it appears to the nodes that they are linked to each other via two dedicated point-to-point connections. A single wavelength can be dropped to an Ethernet switch connected to CWDM APD GBICs. Services are distributed throughout the building via CAT5 twisted-pair cables. Most Ethernet switches are equipped with 24 or 48 CAT5 unshielded twisted-pair 10/100 Ethernet ports. The GBIC slots inside the switches were originally intended for building the backbone inside a building, but you can also use them to create a backbone throughout a small city. The additional 8 dB of link budget in an APD comes at a relatively small price. Today, these diodes, which require a rather high bias for operation, are considered exotic, but they represent a simple way to enhance metropolitan area networking because the extra link budget can be used for WDM and switching. The higher sensitivity in an APD receiver also has an economic advantage. The cost of optical amplification is approximately $1,000 per decibel. If an APD receiver at 2.5 Gbps can add 8 dB to a design engineer's power budget for only a few hundred dollars more than a PIN receiver, then the value of an APD solution is obvious. The wide dynamic range of the APD receiver also offers high flexibility.[1]

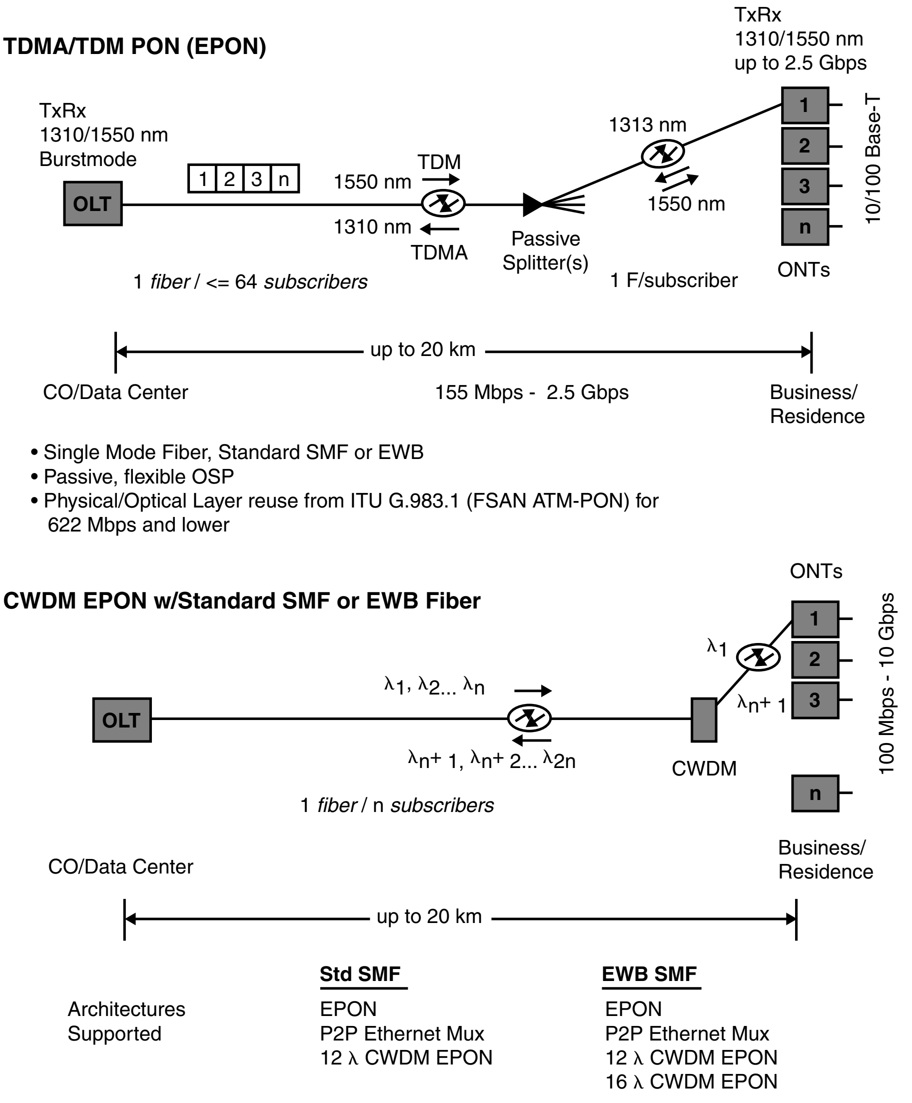


Figure 6-10
TDMA/TDM PON (EPON).

6.4 Other Technologies

Some people advance the concept of bidirectional WDM (BWDM), which we describe briefly in this section to provide an inclusive view of the commercially available technologies.[2] The point of departure of this discussion is the observation that the penetration level of WDM-based solutions to the metropolitan transport market has not been as dominant as in the long-haul market. This is mainly the result of high equipment prices and the lack of integrated manageable solutions with higher-layer technologies such as Packet over SONET (PoS) and the other protocols such as the Spatial Reuse Protocol (SRP). Therefore, initial optical broadband metro applications could deploy CWDM or BWDM, which present a more cost-effective option for metropolitan area markets.

BWDM is a solution based on the transmission of the 1,310- and 1,550-nm wavelengths through a single fiber in opposite directions (see Figure 6-11). With a transmission range of 40 to 120 km (25 to 75 miles), BWDM technology could be used to deliver resilient and cost-effective solutions for metropolitan area deployment. The advantages of BWDM are clear:

- A 100 percent improvement in fiber utilization.
- Simplified fiber connection.
- No resilience penalties. Intelligent protection switching (IPS), enhanced-IPS (E-IPS), and automatic protection switching (APS) are fully supported.
- Cost-effective.
- A range of 40 to 120 km (25 to 75 miles), which is suitable for metro deployment.

6.4.1 Cisco Implementation of BWDM

Cisco equipment (for example, ONS 15190) integrates the BWDM technology within their SRP and PoS technologies. The BWDM line card consists of two single SC ports, each fiber transmitting in both directions. The two wavelengths are converted to an electronic signal on the line card and are placed on the internal bus of the equipment. The traffic is then switched to the appropriate standard line card, which in turn is connected to a PoS or SRP device (see Figure 6-12).

BWDM offers advantages to both SRP rings and PoS links. In metropolitan area environments, BWDM enables the deployment of SRP rings, requir-

Figure 6-11
BWDM.

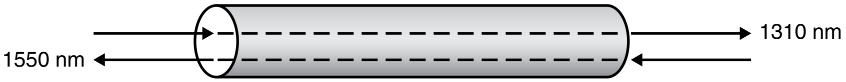
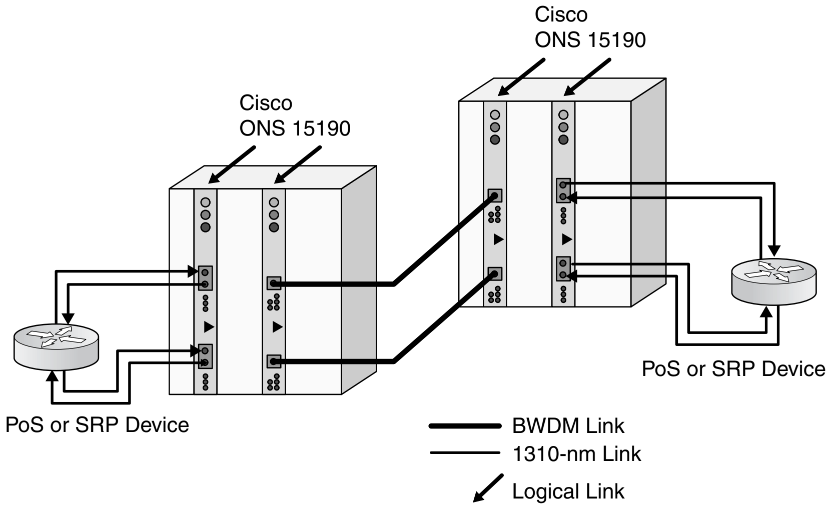


Figure 6-12 BWDM connections to the Cisco ONS 15190.



ing half the fiber required by conventional rings. Because the transmission is in opposite directions, BWDM rings are viewed to be as resilient to resource losses (such as fiber cuts or node failure) as conventional rings. The following section describes various scenarios detailing the events following a loss of resource (fiber cut). In the first case, events are described for conventional rings; in the second case, the events are for BWDM-based rings.

Single- and Dual-Fiber Cuts in the SRP Ring SRP technology is capable of recovering from either a single- or dual-fiber cut. If a single fiber is cut in the SRP ring between nodes A and B, the following occurs:

1. Node B detects the fiber cut and then wraps the ring.
2. Node B sends a short message to node A on the inner ring and a long message through nodes C and D on the outer ring.
3. On receiving the short message from node B, node A wraps the ring and sends a short message to node B (this message is lost because of the fiber cut) and a long message through nodes D and C.

The result of these operations is that the ring is now wrapped and all nodes are connected. (See Figure 6-13.)

If a dual fiber is cut in the SRP ring between nodes A and B, the result is exactly the same as for the single-fiber cut. The following occurs:

1. Node B detects the signal failure and wraps the ring.
2. Node B then sends a short message to node A on the inner ring and a long message through nodes C and D on the outer ring.

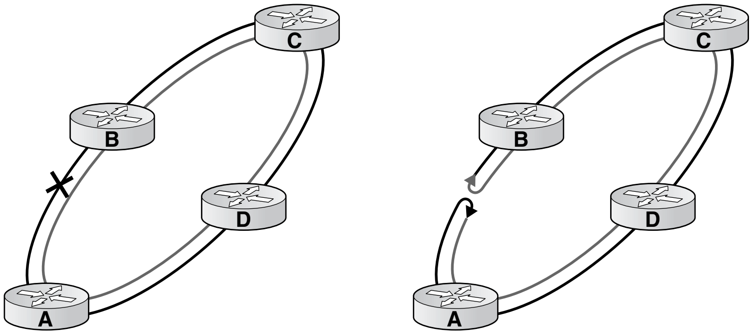


Figure 6-13
IPS solution to single-fiber cut.

3. Node A also detects the signal failure and wraps the ring.
4. Node A then sends a short message to node B on the outer ring and a long message through nodes D and C on the inner ring.

Fiber Cut in BWDM Fiber A single-fiber cut in a fiber-optic cable carrying BWDM signals is actually analogous to a dual-fiber cut because both the receive and transmit signals are disrupted. However, a system will recover from this situation as in the previous scenario of a single-fiber cut. (See Figure 6-14.)

In this example, nodes D and E are connected to the ring via a BWDM link. If a fiber is cut on one of the BWDM links, the following occurs:

1. Node D detects the signal failure and wraps the ring.
2. Node C also detects the signal failure and wraps the ring.

The result of these operations is that the ring is now wrapped and all nodes are connected.

6.4.2 BWDM Applications

In a wide area network (WAN) environment with limited fiber availability, BWDM presents a solution for SRP ring deployment or PoS connectivity. In Figure 6-15, BWDM technology enables an SRP ring to be deployed using the existing point-to-point infrastructure (between locations A and B and between C and D). As shown previously, BWDM-based SRP rings are as resilient as traditional SRP rings. Furthermore, using the BWDM-equipped concentrator to deploy SRP rings provides many advantages such as E-IPS, ring management, maintenance with no service interruption, and more simplified deployment of the ring. (See Figure 6-15.)

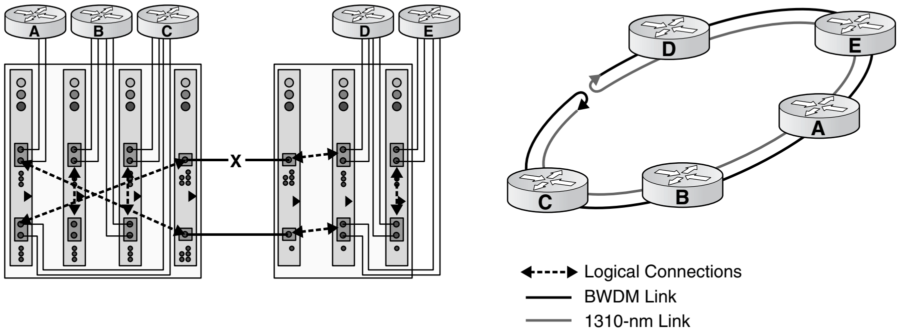


Figure 6-14
Cut in BWDM fiber.

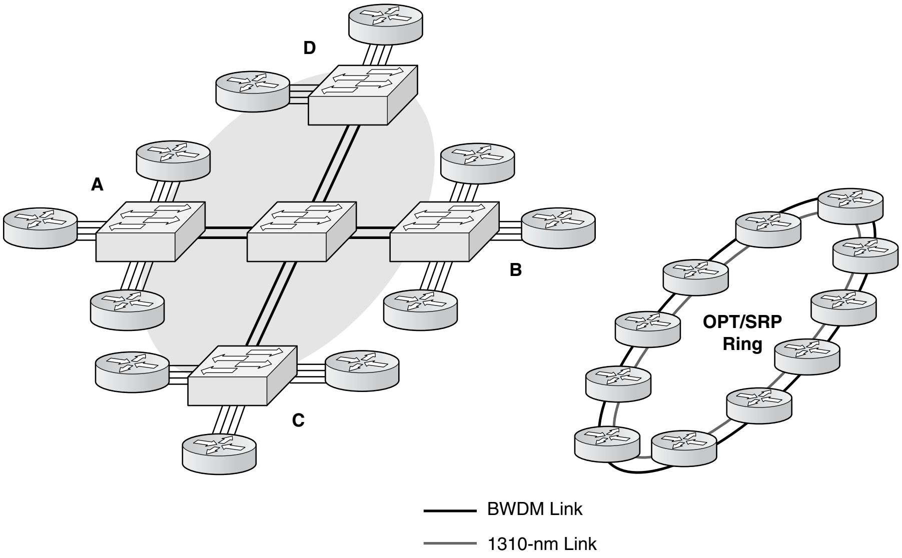


Figure 6-15
BWDM deployment in SRP rings.

In Figure 6-16, BWDM technology is used to enhance PoS point-to-point connections to support SONET/SDH APS protection. Using a single pair of fibers, the bidirectional traffic of the working link can run over the first fiber, enabling the second fiber to act as the protection link. Alternatively, a mesh connection topology can be supported between the four routers using the two fiber pairs.

BWDM technology is also usable for deploying SRP rings in the metropolitan area. In Figure 6-17, all the routers are part of the same SRP ring. By employing BWDM technology, each of the two systems may be connected

Figure 6-16
BWDM deployment in PoS mesh.

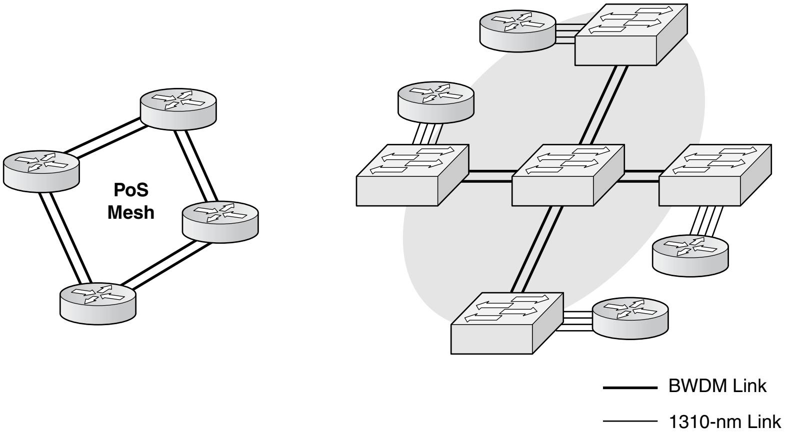
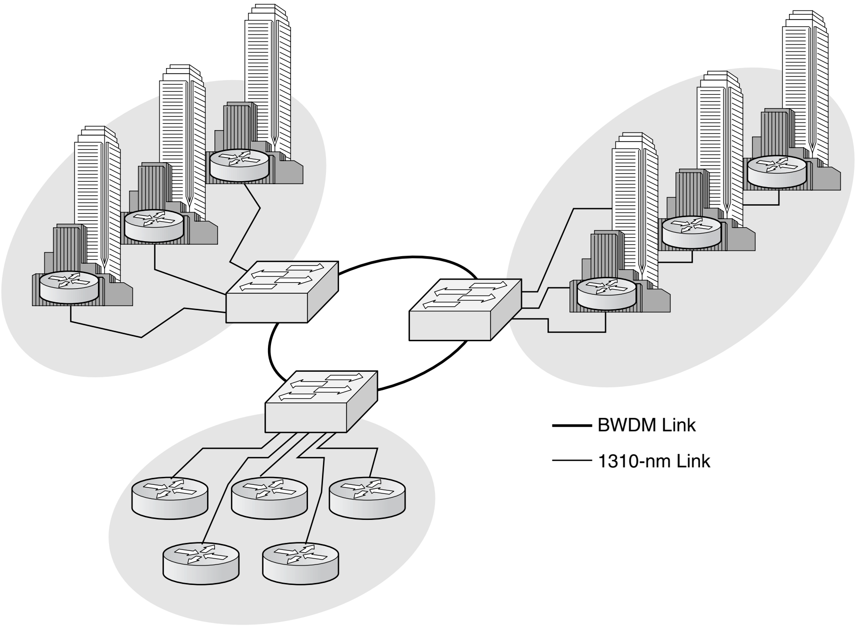


Figure 6-17
BWDM deployment in MAN.



via a single optical fiber and still maintain all SRP ring functionality. Such deployment also provides an enhanced resiliency capability (E-IPS), which is very important when deploying such an SRP ring in the metropolitan area, so that the failure of an individual node will not affect the traffic between the other routers on the ring.

To wrap up the chapter, the main attributes of each type of WDM previously discussed are summarized here:

- CWDM uses a much wider wavelength range (1,200 to 1,600 nm) with a minimum 20-nm wavelength gap between any two channels. Optical components used for CWDM are less accurate and thus less expensive because of the 20-nm wavelength gap.
- DWDM refers to the capability to carry a large number of channels, where large may be 100 channels or more. DWDM equipment usually uses the 1,530- to 1,565-nm wavelength range with very high laser accuracy requiring temperature-controlled components. The standards today define a 0.8-nm wavelength gap between any two channels, with some manufacturers demonstrating equipment with 0.4-nm and even 0.2-nm accuracy.
- BWDM refers to the capability to carry two channels in different directions on a single fiber. Two distinct wavelengths may be used in BWDM technology: 1,310 and 1,550 nm.

End Notes

1. J. Christopher Dries, "InGaAs Avalanche Photodiodes Enable Low-Cost CWDM-Based Metropolitan Area Networks," Cahners ECN Cahners Business Information, www.cahners.com.

2. Cisco White Paper, "WDM-Based Metropolitan-Area Deployment of SRP and PoS with the Cisco ONS 15190," 2001.

3. This section is based in its majority on the white paper by Marcus Nebeling, "CWDM: Lower Cost for More Capacity in the Short Haul," Fiber Network Engineering, Inc., Livermore, Calif., www.fn-eng.com.

4. This illustrates that it is rather challenging to make any money with this specific TLS service; typically, carriers have supported 10 Mbps on ATM, not optics.

5. John George, "Optical Architectures and Fibers," IEEE EFM Meeting, Hilton Head, N.C., March 12–15, 2001.

6. R. Michalzik, G. Giaretta, A.J. Ritger, and Q.L. Williams, "10 Gb/s VCSEL-Based Data Transmission over 1.6 km of New-Generation 850-nm Multimode Fiber," IEEE Lasers and Electro-Optical Society (LEOS)

Annual Meeting, 1999, Postdeadline Paper PD1.6, San Francisco, Calif., November 1999.

7. G. Giaretta, R. Michalzik, and A.J. Ritger, “Long-Distance (2.8 km), Short-Wavelength (0.85 μm) Data Transmission at 10 Gb/sec over New-Generation High-Bandwidth Multimode Fiber,” Conference on Lasers and Electro-Optics (CLEO) 2000, Postdeadline Paper CPD13, San Francisco, Calif., May 2000.

8. R.V. Penty, M. Webster, A.B. Massara, and I.H. White, “Physical Layer Strategies for 10 Gigabit Ethernet,” in Proceedings of 50th Electronic Computer and Technology Conference (ECTC), Las Vegas, Nev., May 2000, 487–490.

9. K. DrÄogemÄuller, D. Kuhl, J. Blank, M. Ehlert, T. Kraeker, J. Häohn, D. Klix, V. Plickert, L. Melchior, I. Schmale, P. Hildebrandt, M. Heinemann, F.P. Schiefelbein, L. Leininger, H.D. Wolf, T. Wipiejewski, and A. Ebberg, “Current Progress of Advanced High-Speed Parallel Optical Links for Computer Clusters and Switching Systems,” in Proceedings of 50th Electronic Computer and Technology Conference (ECTC), Las Vegas, Nev., May 2000, 1,227–1,235.

10. R. Michalzik, G. Giaretta, K.W. Goossen, J.A. Walker, and M.C. Nuss, “40 Gb/s Coarse WDM Data Transmission with 825-nm Wavelength VCSELs over 310 m of High-Performance Multimode Fiber,” in Proceedings of the 26th European Conference on Optical Communication (ECOC) 4, Munich, Germany, September 2000, 33–34.

11. Rainer Michalzik, “Four-Channel Coarse WDM 40 Gb/s Transmission of Short-Wavelength VCSEL Signals over High-Bandwidth Silica Multi-Mode Fiber,” *Annual Report 2000*, Optoelectronics Department, University of Ulm, Ulm, Germany.

CHAPTER 7

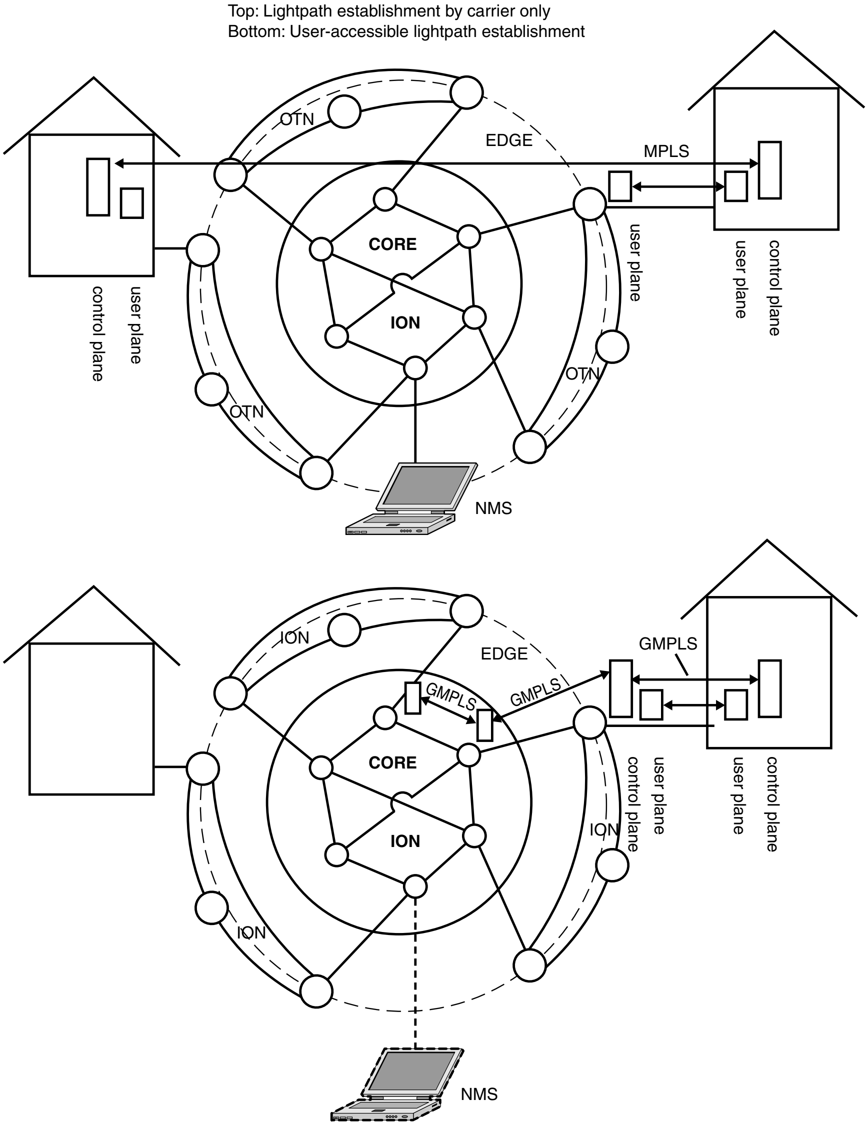
All-Optical Networks

Let's take a critical look at the evolving optical technologies that may see carrier deployment in the next two to five years. They include all-optical networks without real-time user signaling control, which are often called *Optical Transport Networks (OTNs)*, and all-optical networks equipped with real-time signaling control, which are often called *Intelligent Optical Networks (IONs)*. Unsurprisingly, carriers lean toward OTNs, whereas technology vendors favor IONs. In part, the vendor position is determined by the fact that ION implementations require more network equipment and Customer Premises Equipment (CPE) than OTN implementations.

Currently, optical communication is mainly limited to dense wave division multiplexing (DWDM) point-to-point systems. All-optical networks are expected to see increased penetration in the next few years. These transparent optical networks (which eliminate optical-to-electrical-to-optical [O-E-O] conversions) offer the cost-effectiveness that is necessary for infrastructures supporting near-term as well as long-term services. However, when Synchronous Optical Network/Synchronous Digital Hierarchy (SONET/SDH) networks are replaced, the optical network needs to provide SONET-like functions such as protection; switching/routing; Operations, Administration, Maintenance, and Provisioning (OAM&P); and performance monitoring.

IONs require a control mechanism to support dynamic session establishment and routing over the underlying optical network. One of the questions of interest in ION scenarios is whether the lightpath establishment must be end to end (thereby affording the customer complete service control) or only in the core of the network (thereby affording the carrier ease of backbone provisioning). These two scenarios appear in Figure 7-1. Control mechanisms have been proposed for adoption from intranet-based routing developments in multiprotocol label switching (MPLS). MPLS is an enabling technology for next-generation networks that was originally conceived to combine the flexibility of Internet Protocol (IP) hop-by-hop routing with the quality of service (QoS) and traffic-engineering benefits of connection-oriented cut-through Layer 2 protocols such as Asynchronous Transfer Mode (ATM). In its generalized manifestation, MPLS is becoming one of the central elements of the ION architecture. The carriers' acceptance of end-to-end user session establishment remains to be demonstrated, however. Previous efforts at signaling for packet services based on the User-Network Interface (UNI), Frame Relay services, and ATM services have all received lackluster market response (and carrier support) in spite of the copious developments in the standards arena. Nearly all Frame Relay and ATM services are based on permanent virtual connections. This would point to a value of ION in the core tier of the overall network as a mechanism for a

Figure 7-1
ION options.


Top: Lightpath establishment by carrier only
Bottom: User-accessible lightpath establishment

carrier's internal traffic engineering and restoration, rather than an end-to-end capability. It is not that the concept is unsound any more than the concepts of International Telecommunication Union Standardization Sector (ITU-T) Q.931/ISDN, Q.933/FR, or Q.2931/ATM are unsound;[1] it is just that accepted cost-effective solutions are historically lacking in the market. The control mechanism per se will be covered in the discussion of Generalized MPLS (GMPLS) in Chapter 8, "GMPLS in Optical Networks."

7.1 Background and Drivers for ION

This section looks at the business and user landscape within which optical networks have to be deployed. Innovation in optical components and transport technologies has been steady over the last few years and the same is true of optical switching. The optical industry has advanced the concept of ION as a control apparatus for the OTN. ION affords a transition from hybrid electro-optical networks to fully optical networks. This technology also supports a dynamic, multiwavelength, mesh-oriented design with a real-time signaling mechanism. Data-oriented handoffs such as Gigabit Ethernet (GbE) and 10 GbE are increasingly becoming an important end-user requirement for intranet connectivity across a wide area network (WAN). ION is an approach to optical networking that enables the delivery of services, such as the ones just mentioned, over lightpaths (*lightpaths* are high-capacity, all-optical channels). Specifically, ION maps services onto waves (lambdas) and provisions, switches, and manages these lightpaths. In ION, the physical layer of the network is optical, not SONET/SDH. Figure 7-2 depicts a layer architecture view of ION. As you look at the figure, and throughout this chapter, you'll find it helpful to keep this formula in mind:

$$ION = OTN + \textit{control plane capabilities}$$

Requirements affecting the carrier space also continue to change. Data now dominates the volume of traffic compared to voice, although voice continues to generate as much as 80 percent of the total revenues for certain carriers. Major networking applications relate to the interconnection of routers for either enterprise or Internet service provider (ISP) scenarios. During the 1990s, the data interconnection requirements were met using traditional private lines such as T1 and T3, which in turn may have been carried on a SONET/SDH infrastructure. In other cases, IP was mapped directly onto SONET (using Packet over SONET [PoS] technologies, as discussed in Chapter 3, "Traditional SONET Architectures"). However, according to some people, existing ring-based optical architectures could prove insufficient for future requirements. At the same time, advances in optical technologies afford new architectural opportunities. There is also the potential for the emergence of requirements for new direct optical services

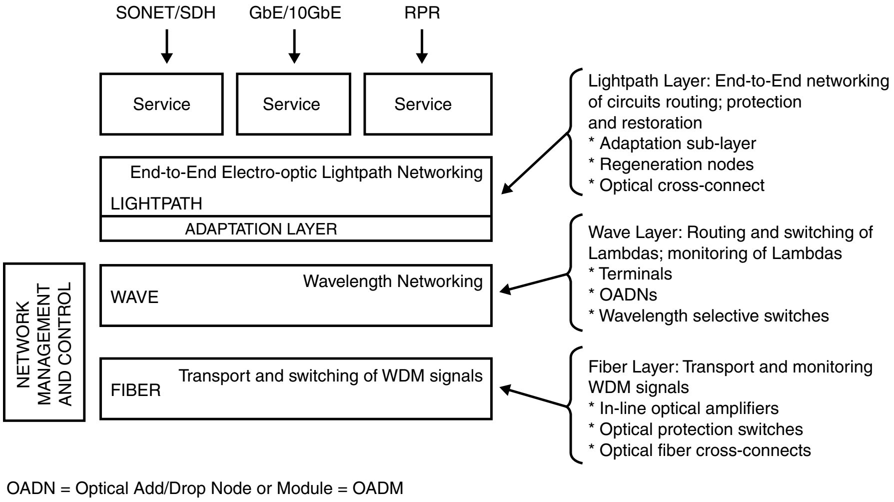


Figure 7-2
ION layers.

(switched lambdas). Naturally, there is also keen interest in reducing service provisioning time, if possible.

In addition to technological changes, observers also anticipate architectural changes. The architecture of evolving optical networks could increasingly be based on mesh configurations rather than cascaded rings. Figure 7-3 depicts a hypothetical transition to a mesh architecture from a ring-based baseline. Keep in mind that the architecture shown on the right is more characteristic of a carrier's physical infrastructure than a user's overlay logic in the United States, where a low-competition scenario means that only a handful of megacarriers provide the end-to-end set of services. The architecture is nearly perfect for the needs of these supercarriers. It is reminiscent of the predivestiture environment or perhaps of the Postal, Telegraph, and Telephone (PTT) model of "one carrier has it all." The more complex architecture shown at the bottom left of Figure 7-3 is in reality more amenable to and representative of a multicarrier, multiplatform, multispace, and multipurpose competitive environment. Mesh networks are often implied by simplified diagrams from equipment vendors and consultants to have abstract appeal, but such presentations fail to take into account the economic and business realities of networking in the United States. Assuming a best-in-breed stance, we can agree that a mesh network is generally more suited to long-haul networks, whereas a cascaded set of rings hanging off this mesh is more suited for metro applications. See Figure 7-4 for an elaboration of this viewpoint.

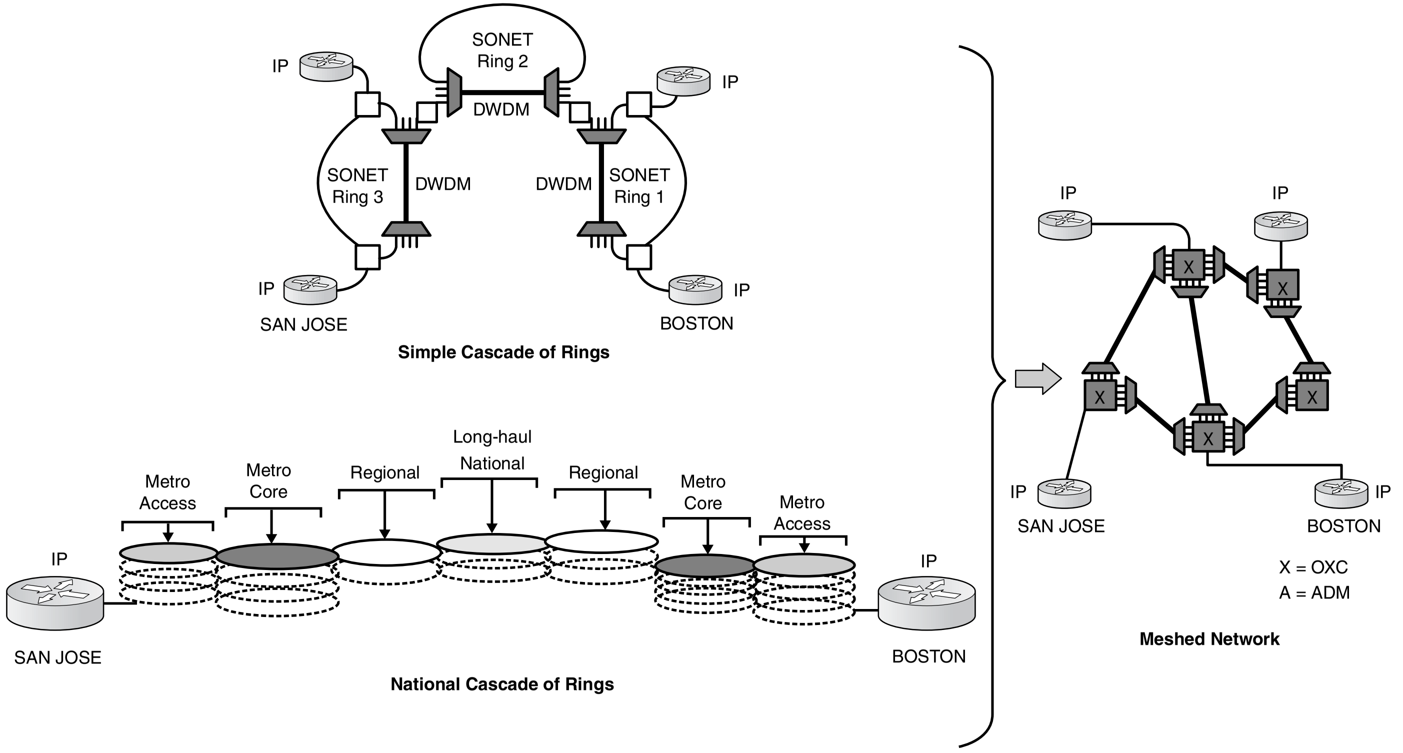
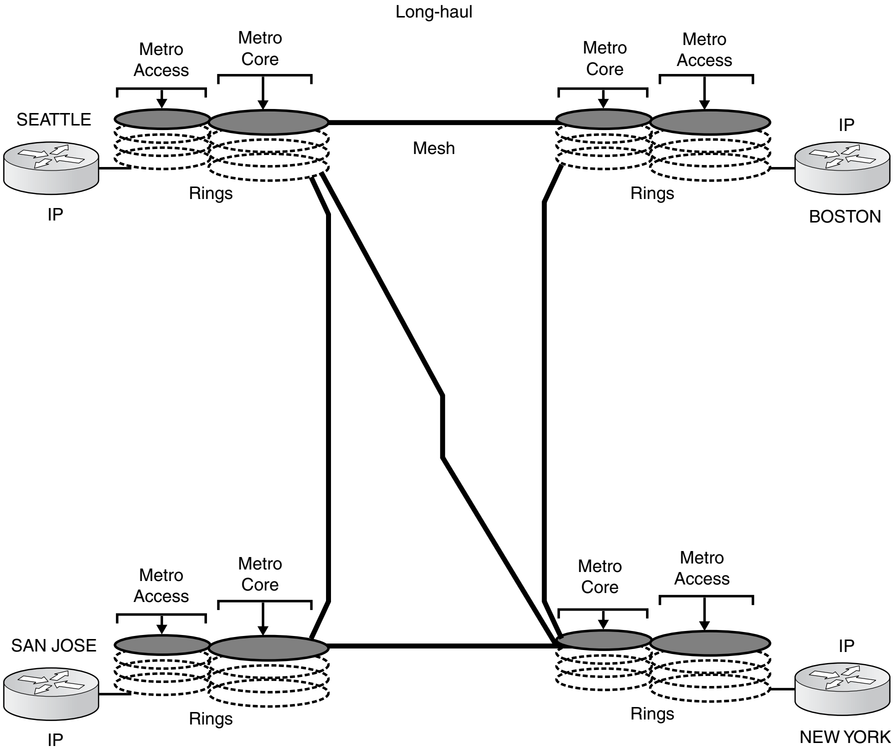


Figure 7-3
Potential transition to meshed networks.

Figure 7-4
Hybrid architectures for the future: mesh long-haul and rings in a metropolitan area network (MAN).



7.1.1 Advantages of ION

Proponents of ION see networks evolving in the next two to five years with the following characteristics:[2]

- IP over optics
- Wave-switched mesh networks
- Switching and managing lightpaths under software control
- Fast provisioning
- Millisecond protection and restoration
- New services, such as new lightpath services, protection on a service level agreement (SLA) metric, value-added optical VPNs (OVPNs), and so on

The motivations for ION include cost-effectiveness, flexibility, and new revenue possibilities. Advocates view ION as a system for bandwidth creation, distribution, and management, with optics being the only physical layer of the network. Whereas DWDM as a technology provides the equivalent of fiber multiplication and moves waves over distances, ION is a system for delivering services over waves. Namely, it moves waves over distances, it maps services onto waves, and it provisions, switches, and manages lightpaths. Switching lightpaths is cheaper than switching individual packets, particularly as the transactions (flows) get larger (longer). The ION design eliminates O-E-O conversions and simplifies the network architecture and the supportive infrastructure. From a flexibility standpoint, if ION were deployed, it would enable just-in-time provisioning of end-to-end lightpaths operating at OC-48, OC-192, or OC-768 speeds. (In this context, *end-to-end* means core network element [NE] to core NE if the technology is only deployed in the core or edge NE to edge NE if the technology is deployed across the entire network.) Traffic engineering would be done at the lightpath level. New revenue opportunities could arise from the fast provisioning of wave services.

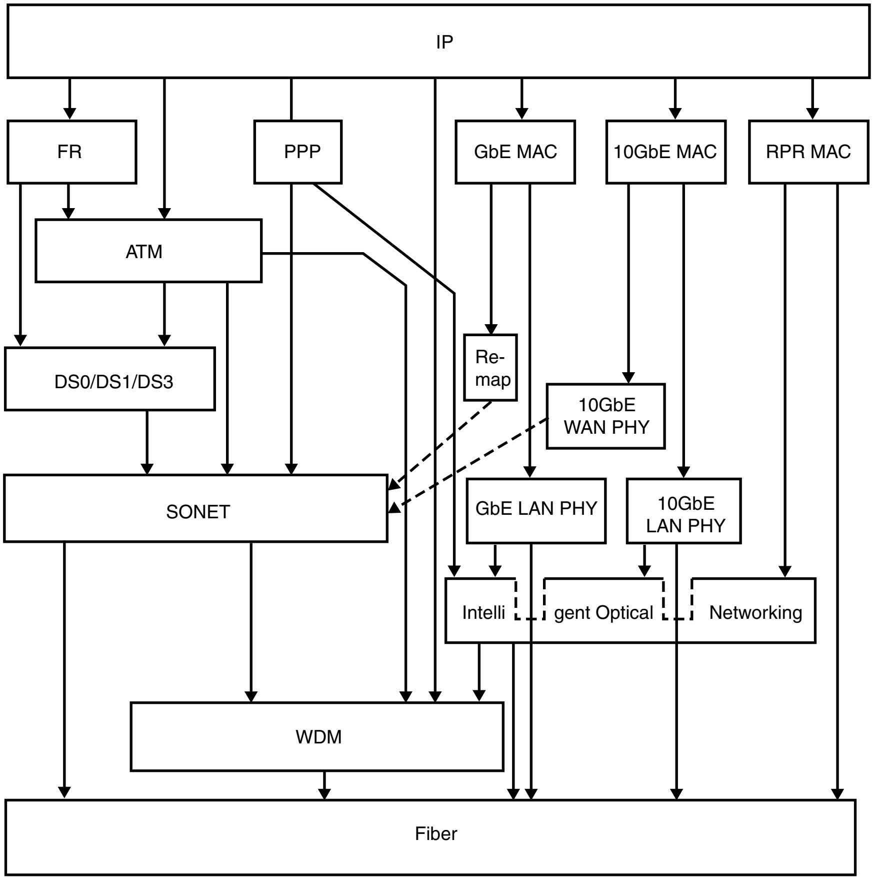
ION, however, also needs to support traditional private lines (for example, Frame Relay services and T1 lines currently represent the staple of data communication networking in the Fortune 5000 companies), SONET-based services, and any Ethernet-based transparent local area network (LAN) services (TLS) that may evolve. As noted, TLS-like services have been around for many years, and it is not clear how much further they will develop in the near future. Since the mid-1980s, TLS has been offered via a Resilient Packet Ring (RPR) type of architecture such as IEEE 802.6 and Bellcore's Switched Multimegabit Data Service (SMDS). It runs over dedicated point-to-point or inverse-multiplexed links, SONET, ATM, and even over GbE. TLS certainly can take credit for being open to and having tried all sorts of technologies over its 15-year existence.

Figure 7-5 depicts a transition from Chapter 1, "Advances and Opportunities in Next-Generation SONET and Other Optical Architectures," from IP to a scenario where IP runs on OTN or even ION. However, this transition assumes that metropolitan area service providers (MASPs) and long-haul service providers will offer affordable lambda services to end users. In 2001, we saw an unfortunate shrinkage in the number of carriers offering alternative, high-end, and innovative services. If carriers don't provide these services, the promise of ION and GMPLS could only be realized over private optical networks, which are cost prohibitive for nearly all end users.

ION presents the carriers with a number of prima facie benefits:

- Switching lambdas is cheaper than switching packets and gives users better QoS and throughput.
- It eliminates costly O-E-O conversions.

Figure 7-5
Evolving protocol stack showing ION.

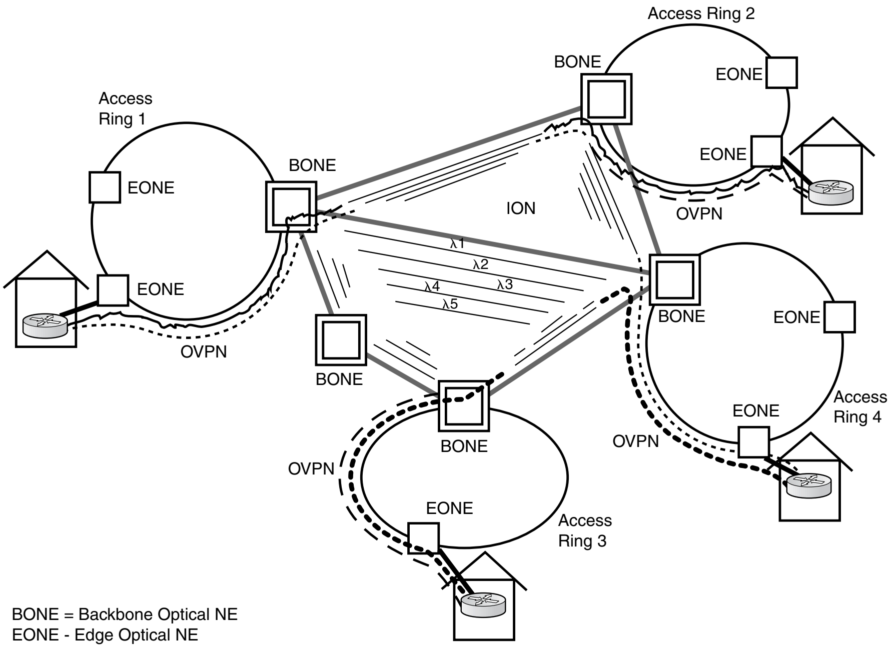


- It cuts down equipment while consolidating systems and network infrastructure.

Furthermore, there may be flexibility and management advantages in good implementations, such as just-in-time provisioning, point-and-click service provisioning, and traffic engineering at the lambda level. For those buildings in the United States with fiber facilities, revenue opportunities reside in fast-provisioned, bandwidth-on-demand lambda services and OVPNs. See Figure 7-6 for a graphical example.

SONET/SDH and Ethernet private-line services can be supported with ION. Architecturally, ION enables the gradual and seamless adoption of new technologies into the network and supports the migration to mesh networks. In contrast to dark fiber and unprotected WDM, it provides access to protection bandwidth. Developers designing ION capabilities have to focus

Figure 7-6
OVPN.



on carrier-class services, however, because that's what customers have tended to demand. The customers want

- Network autodiscovery
- Availability
- Reliability
- Fault isolation and performance monitoring
- Customer Network Management (CNM) capabilities

7.1.2 Issues Impacting ION

On the negative side of the ION discussion, carriers would lose some control if they allow users to fire up and tear down lightpaths in real time all over the network. There could also be a loss in revenue if users only fire up lightpaths for a total of a few minutes a day, instead of paying for the service on a dedicated basis. Furthermore, mesh networks might need to rely on or lead to what people have called a *god box*, a does-it-all device that merges an add/drop multiplexer (ADM), a digital crossconnect system (DCS), and other functions. Although such an approach could eliminate ring-to-ring interconnects and reduce 20 to 40 percent of the optical trunks, the reality is that according to industry estimates, there already are 100,000 SONET rings deployed in the United States, and there is no motivation (financial or otherwise) to decommission any such equipment or fiber links. Naturally, a greenfield network could be designed in a different fashion; however, no such greenfield opportunities exist in industrialized countries at a time of fiber glut in segments of the network. Although you can have OTNs and IONs without mesh architectures, proponents state that meshes are beneficial.[3]

To summarize this discussion on mesh topologies, a 10- to 15-node super-backbone mesh spanning the United States might have an architectural appeal over cascaded rings. However, the problem is that proponents compare the alternatives of a new mesh network versus a new cascaded-ring network, instead of what they should compare, which is a new mesh network versus a cascaded-ring network that was already paid for, deployed, engineered, tested, and in service. Further, the argument could be made that provisioning service over, say, three rings could take five provisioning steps, whereas a state-of-the-art mesh network could support automatic flow-through provisioning from a single workstation command that identifies the two endpoints, thereby reducing operational overhead. This implies that a technician could provision more orders in a day—assuming the demand for an additional order existed—or the carrier could reduce the provisioning staff. Again, the proper comparison is an OAM&P on a new mesh network versus an OAM&P on a cascaded-ring network that was already paid for, deployed, engineered, tested, in service, and operations support system (OSS) integrated (and had a trained staff). We are not making the case for slow innovation; we are just urging a proper financial analysis of the two alternatives.

Equipment vendors and consultants continue to make the case that the volume of data traffic keeps increasing over time, while the volume of voice traffic tends to remain stable over time. This would drive new architectures, presumably forcing a transition away from SONET/SDH systems. The following paragraph attempts to bring some perspective to this issue.

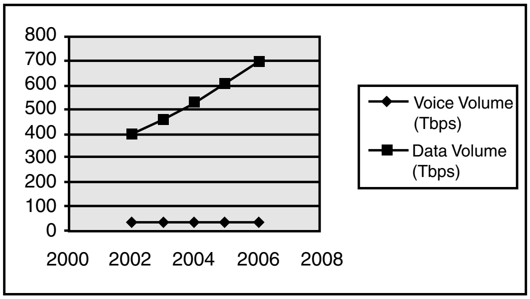
The research firm Ovum places the voice volume at about 30 Tbps and data at about 400 Tbps in 2002.[4] By 2006, the volume of voice will be approximately the same, while the volume for data is expected to grow to 700 Tbps (this equates to a compound annual growth rate [CAGR] of 15 percent). Now take into account numbers published by the New Paradigm Group, which tend to show (at least for the United States) a revenue parity for data and voice for Competitive Local Exchange Carriers (CLECs) in a few years. Some people quote a more slanted 80 to 20 percent split with voice being the 80 percent. Our critical take-away from the vendors' assertions regarding data versus voice volume is that the revenue per bit for data services is one-thirteenth the revenue per bit for voice services (see Figure 7-7). Although this configuration may need tweaking in later years, our figure points to the fact that revenue on data channels is much smaller

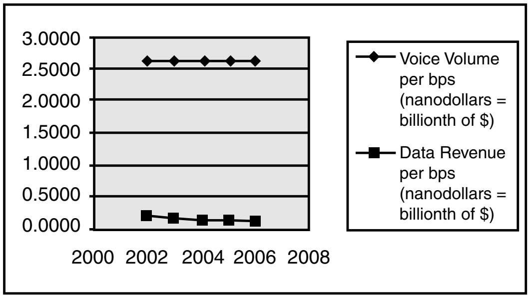
Figure 7-7
Voice/data volume versus voice revenue per bit.

	Voice Volume (Tbps)	Data Volume (Tbps)
2002	30	400
2003	30	460
2004	30	529
2005	30	608
2006	30	700

	Voice Volume per bps (nanodollars = billionth of $)	Total Revenue Ratio d-t-v	Data Revenue per bps (nanodollars = billionth of $)
2002	2.6042	0.50	0.1953
2003	2.6042	0.50	0.1698
2004	2.6042	0.50	0.1477
2005	2.6042	0.50	0.1285
2006	2.6042	0.50	0.1116

Note: Assumes voice calls
are charged at $0.01/minute





than revenue on voice channels. As a simplified convergence point, in a few years every U.S. household could subscribe to a $25 per month Internet service (for an approximate yearly data revenue of $300×100,000,000=$30 billion), a $25 per month access line fee (for an approximate yearly access revenue of $300×100,000,000=$30 billion), a $25 per month postalized all-you-can-use long-distance voice fee (for an approximate yearly voice revenue of $300×100,000,000=$30 billion), a $25 per month all-you-can-use wireless fee (for an approximate yearly wireless revenue of $300 × 100,000,000 = $30 billion), and a $25 per month cable TV fee (for an approximate yearly cable TV revenue of $300 × 100,000,000 = $30 billion). That amounts to $150 billion a year for residential customers.

On the commercial side, some carriers see a growing demand to interconnect OC-48/OC-192-based routers and an emerging OC-768 requirement—demand for GbE and 10 GbE TLS services. They also see that rings are insufficient for large pipes, the interconnection of cascaded rings does not support end-to-end provisioning and management, and IP does not always need the restoration protection of rings. However, end-to-end provisioning is related to the availability, or lack thereof, of fiber facilities in the majority of customer buildings.

Traditionally, provisioning in optical networks has required manual planning and configuration, resulting in setup times of days or even weeks. Can autoprovisioning help? Any individual who has ever worked in a provisioning department in a carrier understands that in nearly all instances, the provisioning time interval is allocable 90 percent to the access facility side of the network and 10 percent to the backbone side. A provisioner could wait days, weeks, or even months for a fiber to be brought into a building from the street, through a riser, or onto a floor, and for power to be installed to run a carrier's NE (such as an edge device, an ADM, and so on). Once the network is in place, you can typically change bandwidth parameters in a matter of a few minutes or hours, even when utilizing traditional technologies.

Although disposing quickly of the misconception that any near-term technology will alleviate provisioning time for optical services in the 4.6 million commercial buildings in the United States that do not have fiber links, you may subscribe to the assertion that all-optical networking requires optical (wavelength) switching, and ION requires wavelength switches that are dynamically reconfigurable (again, the switching may only be automated in the core). Notable recent efforts include MPLS and MPλS, which is now known as GMPLS.

The introduction of WDM technology represents a step toward the deployment of all-optical networks. Optical add/drop modules (OADMs) (also called *optical add/drop nodes*) and optical crossconnects (OXCs) are the additional steps taken. The migration would be completed by the addition of cost-effective and reliable optical switching. OXCs facilitate optical networking in an OTN environment in the sense that lightpaths can be set up end to end (see Table 7-1).[5] However, most lightpaths still have to be provisioned manually. ION proponents argue that optical switches and appropriate signaling protocols could automate this process. However, signaled services have historically failed to see deployment. X.25 supported both provisioned (PVC) and signaled (SVC) connections; only the PVC service has seen widespread deployment. Frame Relay supports both provisioned (PVC) and signaled (SVC) connections; only the PVC service has seen widespread deployment and market interest. ATM supports both provisioned (PVC) and signaled (SVC) connections; only the PVC service has seen deployment of any scale.[6]

7.1.3 Mechanisms to Support ION

In order to be able to enjoy dynamic bandwidth allocation (DBA), particularly in the core where you want to control traffic management, a control mechanism must be deployed. Historically, optical components are manually configured and require a certain amount of time to affect a network-impacting reconfiguration; new control mechanisms (both in the transport plane and control plane) are needed. In addition, it is also important that all

Table 7-1

Comparison of Routers and OXCs

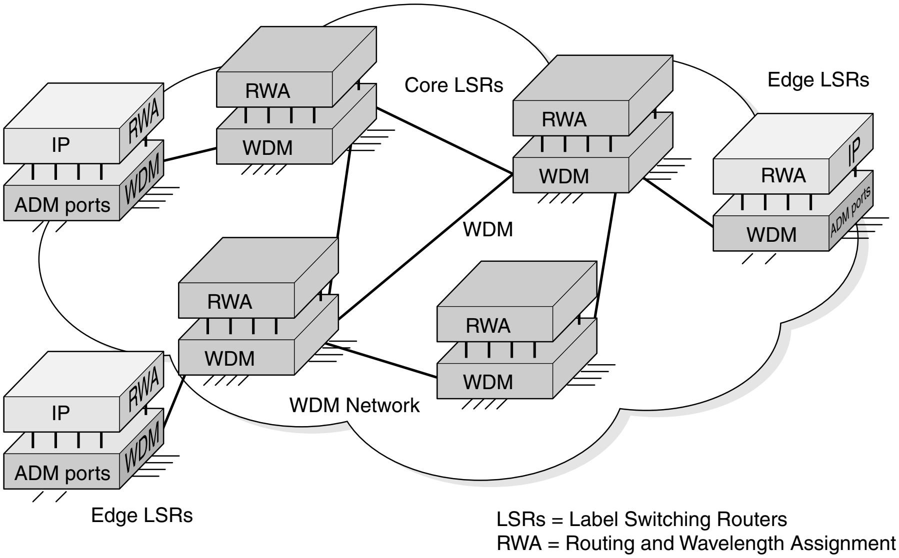
IP Router/Label SW/Router	OXC/Optical Switch
■ Forwards packets based on IP or label header	■ Switches optical signal from the input port to output port
■ Recognizes packet boundaries	■ Does not recognize packet boundaries
■ Electrical or optical interfaces	■ Optical interfaces only
■ O-E and E-O conversions for each packet entering/exiting the system	■ May perform O-E-O or O-O-O for each signal entering/exiting the system
■ IP control plane	■ Control plane for dynamic reconfiguration of optical switch fabric
■ Many sizes and speeds ranging from small metro access up to terabit	■ OC-48, OC-192, and OC-768 framed signals

the optical network elements involved be integrated into an end-to-end management system that facilitates full OAM&P functionality.

The MPLS model developed by the Internet Engineering Task Force (IETF) can provide the signaling mechanism for IONs when applied to an optical network. MPLS is now a subset of GMPLS—one candidate to control the OXCs and optical switches. A theoretical advantage to GMPLS is its capability to work with a single integrated control plane, controlling both the IP and the OTN in an IP over OTN network scenario, which basically equates to the ION. Figure 7-8 shows this scenario. In such a wavelength-routed optical network, lightpaths are set up between endpoints to meet real-time connectivity requirements. Under this dynamic connectivity environment, a lightpath is taken down after a defined period of time (known as the *connection-holding time*). The most important problem in wavelength-routed networks is the routing and wavelength assignment (RWA) problem, which, however, has been proven to be mathematically NP-complete.[7] The multicast RWA problem is even harder to solve than the unicast RWA problem.

Optical switches or switching matrices must have the proper optical characteristics such as low insertion loss, high isolation, and low back reflection. In addition, they need to be reliable, fast, compact, low power consuming, and scalable. As discussed in Chapter 2, "Optics Primer," technologies that are suitable for optical-switching matrices include, but are not

Figure 7-8
IP over a lambda-switched WDM network.

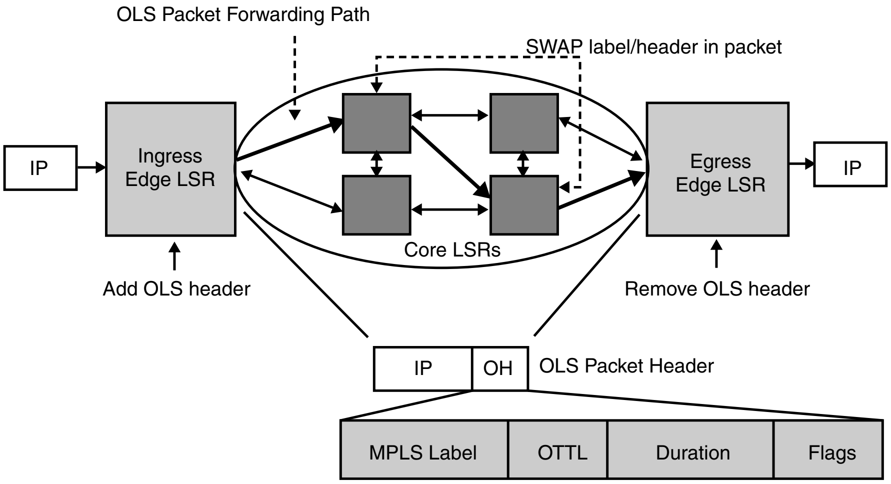


restricted to, Micro-Electro Mechanical Systems (MEMS), some MEMS-like technologies (such as thermocapillary and piezo-electric), liquid crystals, bubbles, thermo- and acousto-optics, and holograms. Beyond technology, there are always issues related to price point affordability. In general, as discussed elsewhere in this book, MAN technology of any kind needs to meet different optimality criteria than that for long-haul applications. For example, while looking at OXCs for metro networks, the cost of crossconnects needs to be much lower. This is because the total link costs in the metropolitan core are not necessarily dominated by fiber costs, which is the case in the long-haul domain. In fact, the opposite may be true for metro access. As an observation, it is not likely that OXCs with wavelength conversion will be used in metro networks in the near future.

One key question is how rapidly the lambda establishment can be set up. If subsecond connections are required, then a full-fledged optical switch is needed. If a carrier is simply trying to support automated fiber provisioning (off of a service order), medium-sized distributed OXCs controlled from a network management station will suffice. Depending on the per-cable fiber count and the number of fibers that are to be flexibly assigned, switching matrix size may start with 64×64. For real-time lambda switching/routing, a full-fledged optical switch will be needed. Wavelength-switching devices (λ switches) enable the switching of each wavelength path independently; dynamic switching is supported via a signaling mechanism. Figure 7-9 illustrates the concept of optical label switching (OLS).

To realize the full promise of OLS, additional technical advancements are needed, including efficient wavelength conversion, optical packet buffer-

Figure 7-9
The concept of OLS.



ing, contention resolution technology, optical reamplify, reshape, and retime (3R) regeneration, and efficient burst-mode receivers. Observers indicate that efficient use of a WDM optical layer requires wavelengths to be viewed as a shared resource—not as individual pipes between edge devices. OLS holds the promise for switching optical packets on a per-packet, per-flow, and per-circuit basis.[8]

A retardant factor to the introduction of ION is that incumbent carriers are not motivated to crosscannibalize lower-speed services that have a high revenue per bit, such as T1s and Frame Relay, with high-speed lambda services.[9] Unfortunately, many new-generation communications carriers who would provide such services have been forced out of business by the negative environment created by the CLECs and Digital Subscriber Line LECs (DLECs) who promised gross margins of 80 percent to venture capitalists and payback (EBITDA positive) in one to two years. When the venture capitalists discovered that these statements were completely untenable, they withheld any and all funding, even to those carriers that always stated more achievable numbers in the 40 percent gross/20 percent net range and who had compelling business cases.

While waiting for the deployment of full ION, we are likely to see less encompassing architectures evolve, such as semiautomated OTNs. The semiautomation applies to provisioning as well as to restoration. For example, optical protection is an important requirement in metro optical networks (MONs), especially for those services that are natively unprotected, such as Fibre Channel and so on. Currently, two-fiber optical uni-/bidirectional path-switched ring (2F-OUPSR/2F-OBPSR) methodologies are typical.[10] The issue is that these protection schemes require a high amount of redundancy. This is because these schemes offer dedicated protection. In order to achieve more efficient protection, techniques that support optical shared protection have to be deployed. A number of concepts exist, including two-fiber shared-lambda protection rings (2F-SλPR) and four-fiber optical bidirectional line-switched rings (4F-OBLSR).[11] Both of these techniques require switch matrices and/or OADMs as basic elements; small switches or OADMs support the capability of *shared* protection. SλPR methods can lead to substantial equipment/fiber savings given that the ring traffic pattern is the nonhubbed type. Figure 7-10 shows the traffic pattern of a fully meshed four-node ring with eight bidirectional channels between each pair of nodes.[12] Table 7-2 compares SλPR and the commonly used OUPSR/OBPSR approach. A total of 32 wavelengths and the traffic pattern from Figure 7-10 are assumed in the table. Optical channel shared protection is viewed by proponents as one of the best solutions for realizing OTN survivability. Standby wavelengths are shared and signals are switched by a wavelength switch.

Figure 7-10
SλPRs.

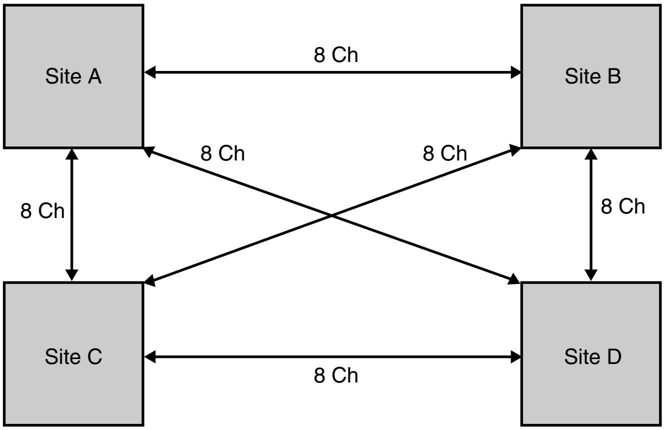


Table 7-2
Comparison Between SλPR and UPSR for a Fully Meshed Four-Node Ring

Ring Systems with 32 Wavelengths	SλPR	UPSR
No. of transponders per ring	176	192
No. of receive transponders per ring	96	192
No. of protection switches per ring	—	96
No. of switch arrays per ring	40	—
No. of fiber pairs necessary	1	2

7.2 Approaches for ION

7.2.1 Control Plane Mechanism

MPLS is a set of protocols developed during the past five years for provisioning and managing core packet networks (see Table 7-3). These protocols were originally developed for intranet applications. MPLS was developed as a packet-based technology and is seeing deployment in core IP-based networks, including converged data and voice networks. MPLS does not replace IP routing, but works alongside existing and future routing technologies to provide high-speed data forwarding between label-switched routers (LSRs) together with the reservation of bandwidth for traffic flows with differing QoS requirements.

Table 7-3

IETF MPLS Documents

Application	IETF Document
Destination-based IP routing	RFC3036
IP VPN	RFC2547
Traffic engineering (RSVP-TE)	*draft-ietf-mpls-rsvp-lsp-tunnel-09.txt*
Traffic engineering (CR-LDP)	*draft-ietf-mpls-cr-ldp-05.txt*
Diffserv traffic engineering	*draft-ietf-tewg-diff-te-reqts-01.txt*
Fast recovery	*draft-ietf-mpls-recovery-frmwrk-03.txt*
MPLS VPN multicast	*draft-rosen-vpn-mcast-02.txt*
Optical	*draft-ietf-mpls-generalized-signaling-06.txt*
L2 transport	*draft-martini-l2circuit-trans-mpls-07.txt*
Interdomain	RFC 3107

Networks that can be controlled by MPLS may be data-centric like those of enterprises or ISPs, voice-centric like those of traditional telecommunications companies, or a converged network that combines voice and data. MPLS overlays a packet-switched IP network to support traffic engineering. Specifically, it enables resources to be reserved and routes to be predetermined. MPLS provides virtual links or tunnels through the network to connect nodes that lie at the edge of the network. For packets injected into the ingress of an established MPLS tunnel, normal IP routing procedures are suspended; instead, the packets are label switched so that they automatically follow the tunnel to its egress.[13]

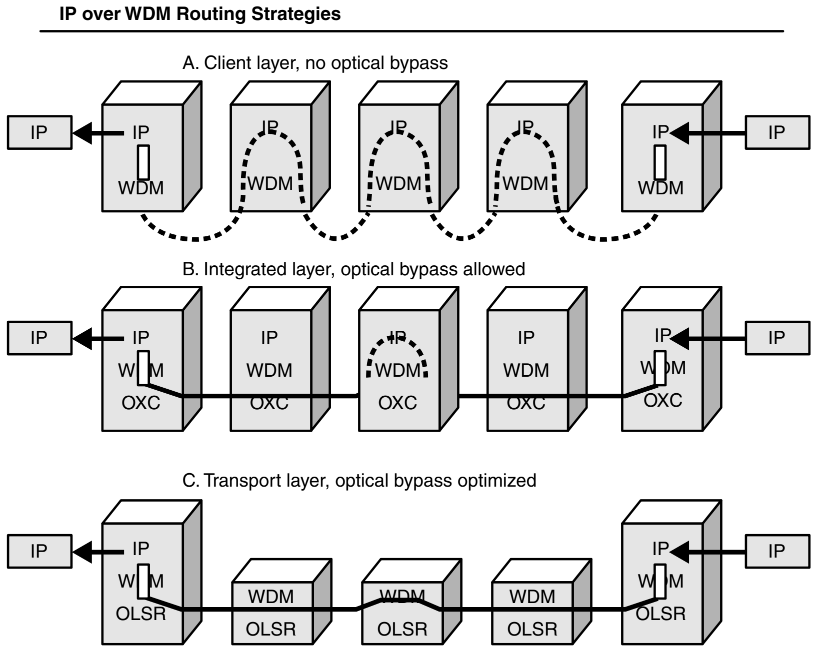
In the core, however, most existing high-bandwidth networks do not currently switch packets. Instead, the bandwidth of the underlying optical fibers is allocated by either frequency-division multiplexing (FDM) (wavelength) or time-division multiplexing (TDM). NEs forward the data by switching particular time slots, wavelengths, or wavebands, or even the contents of entire fibers. This deterministic bandwidth allocation supports the guaranteed throughput for data and targeted QoS. The drawback of this approach is the overhead of the control traffic needed to set up and maintain bandwidth allocations, and, for bursty streams, the potential waste when an application does not use all the bandwidth that was reserved. This last observation, however, needs to be tempered by two fundamental facts that are often ignored by proponents of packet technologies: (1) bandwidth (at least in the long haul) has become a (relatively) inexpensive commodity, and (2) backbone traffic in a moderately to highly loaded network is not bursty when it contains traffic from thousands of sources and is processed by a chain of routers via speed-limited output ports (with deterministic service time). A simple application of the Central Limit Theorem on traffic flows comprised of not trivial (Web-, FTP-, and videostreaming-based) transactions will prove this assertion. In lieu of a Central Limit Theorem argument, researchers can take measurements in an actual moderately to highly loaded network and use the extensive body of techniques from the field of statistics known as *goodness of fit* and determine the distribution of the inter-router traffic.[14] These points are not to be taken as a negative assessment of packet technology, but simply that these factors have to be properly taken into account when designing or pricing out backbone technologies and architectures, and pitching these technologies against technologies that are already deployed.

MPLS proponents counter that where control protocols have been deployed to provision optical networks, they have been proprietary and have suffered from interoperability problems. With the modest success of MPLS in packet-switched IP networks, optical network providers have developed a process to generalize the applicability of MPLS to cover optical networks as well, the result of which is the set of Internet Drafts that are collectively referred to as GMPLS. These drafts extend the following:[13]

- The MPLS data forwarding model, such that it includes current practice in optical networks
- The MPLS control protocols, so that they can be used as a standardized and interoperable way of provisioning optical networks

Figure 7-11 lays out three approaches to optical routing. The top case depicts a fairly traditional approach. This approach is resource intensive because packet processing is needed at every node. The middle case shows an improved architecture, where optical bypass is supported at intermediary nodes. The bottom case enables edge-to-edge optical routing.[8] Some people have called this technology optical label switching (OLS).[8] As mentioned earlier, OLS enables packet switching in the optical domain. Packet forwarding is based on an optical header. The header is a subcarrier multiplexed with the optical data. The label field in the optical header determines packet forwarding. Data is delayed (in a fiber loop) while the header is examined; however, the data never leaves the optical domain. A control function will erase and reinsert the label in the optical header (for example, using an optical notch filter).

Figure 7-11 *IP processing scenarios for an optical network.*



OLSR = Optical Label Switching Routing NE

GMPLS protocols are analyzed in Chapter 8. Related work to standardize the management and configuration of optical networks is ongoing in the development of the Link Management Protocol (LMP) and optical extensions to Open Shortest Path First (OSPF). We will analyze the use of MPLS-based techniques for optical networks, but want to remind you once again that this technology is not a panacea for instant provisioning, particularly when you take the perspective of the end user.

7.2.2 Design Considerations

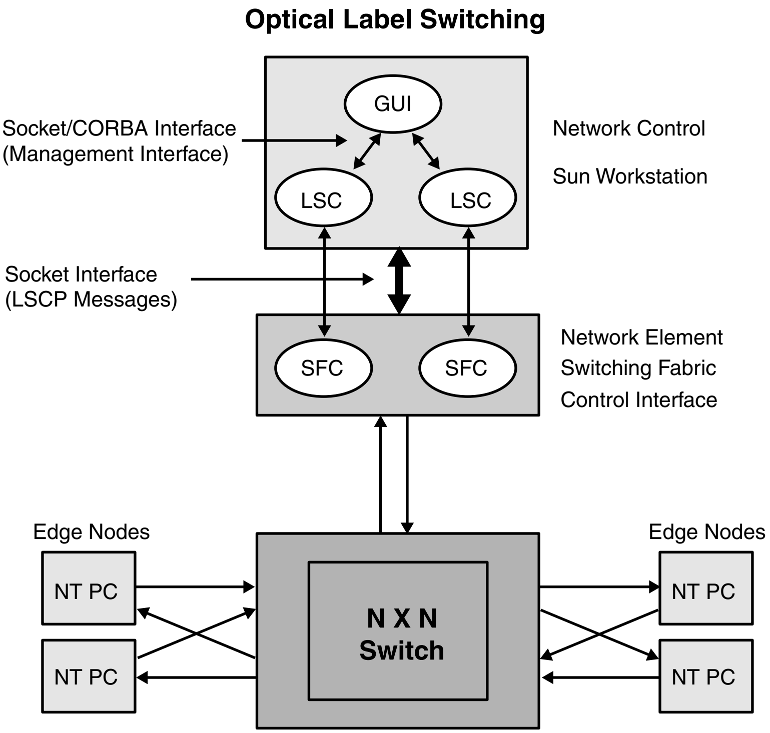
OLS network design goals typically include the following:

- Flexible, scalable, and dynamically reconfigurable optical network
    - Must not mandate that every node must have an IP stack or exclusively an overlay or peer architecture
    - Must scale to accommodate large OLS (WDM) core networks

- Enable the reuse of standard control plane protocols
  - Routing and OSPF, signaling and RSVP, and the switching control interface and LSCP
  - Paths set up and tear down according to MPLS label
  - Integrated layer (IP/WDM) MPλS label-switched path (LSP)

Figure 7-12 depicts a sample conceptual design.[8]

Figure 7-12 Example of OLS node. Source: Gee-Kung Chang, Optical Internet Research, Telcordia Technologies.



- LSC: Label Switch Controller
- SFC: Switching Fabric Controller

Total Decision and Switching time is about 100 ns:

40 ns to acquire destination oriented label and priority information
20 ns to arbitrate control of the switch and output a new switch state
10 ns for $LiNbO_3$ switch and 40 ns for the driver circuit

7.3 A Review of MPLS

This section[15] encapsulates the basic concepts of MPLS. The benefits of MPLS include the following:

- It has the capability to set performance characteristics for a class of traffic.
- It enables VPNs. Using MPLS, service providers can create IP tunnels throughout their network, without the need for encryption or end-user applications.
- It eliminates multiple layers. Most carriers employ an overlay model where ATM is used at Layer 2 and IP is used at Layer 3. By using MPLS, carriers can migrate many functions of the ATM control plane to Layer 3, simplifying network management and network complexity. Carrier networks may be able to migrate away from ATM altogether, which means eliminating ATM in carrying IP traffic.

The following list identifies important Internet Drafts supporting MPLS. The Internet Drafts that are now on the IETF standard track as of July 20, 2001 are as follows:

- *RFC 2702—Requirements for Traffic Engineering over MPLS* Identifies the functional capabilities required to implement policies that facilitate efficient and reliable network operations in an MPLS domain. These capabilities can be used to optimize the utilization of network resources and to enhance traffic-oriented performance characteristics.
- *RFC 3031—Multiprotocol Label Switching Architecture* Specifies the architecture of MPLS.
- *RFC 3032—MPLS Label Stack Encoding* Specifies the encoding to be used by an LSR in order to transmit labeled packets on Point-to-Point Protocol (PPP) data links, on LAN data links, and possibly on other data links. Also specifies the rules and procedures for processing the various fields of the label stack encoding.
- *RFC 3034—Use of Label Switching on Frame Relay Networks Specification* Defines the model and generic mechanisms for MPLS on Frame Relay networks. Extends and clarifies portions of the MPLS architecture and the Label Distribution Protocol (LDP) relative to Frame Relay networks.

- *RFC 3035—MPLS using LDP and ATM VC Switching* Specifies in detail which procedures to use when distributing labels to or from ATM-LSRs when those labels represent forwarding equivalence classes (FECs) for which the routes are determined on a hop-by-hop basis by network layer routing algorithms. Also specifies the MPLS encapsulation to be used when sending labeled packets to or from ATM-LSRs.
- *RFC 3036—LDP Specification* Defines a set of procedures called LDP by which LSRs distribute labels to support MPLS forwarding along normally routed paths.
- *RFC 3037—LDP Applicability* Describes the applicability of LDP by which LSRs distribute labels to support MPLS forwarding along normally routed paths.
- *RFC 3038—VCID Notification over ATM Link for LDP* Specifies the procedures for the communication of virtual channel ID (VCID) values between neighboring ATM-LSRs.
- *RFC 3033—The Assignment of the Information Field and Protocol Identifier in the Q.2941 Generic Identifier and Q.2957 User-to-User Signaling for the Internet Protocol* Specifies the assignment of the information field and protocol identifier in the Q.2941 generic identifier and Q.2957 user-to-user signaling for the Internet protocol.
- *RFC 3107—Carrying Label Information in BGP-4* Specifies the way in which the label-mapping information for a particular route is piggybacked in the same Border Gateway Protocol (BGP) update message that is used to distribute the route itself. When BGP is used to distribute a particular route, it can be also be used to distribute an MPLS label that is mapped to that route.

7.3.1 MPLS Elements

MPLS utilizes label switching to forward data through the network. A label is a small fixed-format field that is inserted in front of each data packet upon entry into the MPLS network. At each hop across the network, the packet is routed based on the value of the incoming interface and label, and dispatched to an outgoing interface with a new label value. In effect, this is similar to virtual channel (VC) switching in ATM. The path that data follows through a network is defined by the transition in label values as the label is swapped at each LSR. Because the mapping between labels is constant at each LSR, the complete path is determined by the initial label value. This is an LSP (see Figure 7-13). A set of packets that should be labeled with the same label value on entry to the MPLS network and that will therefore follow the same LSP is a FEC.[16] Figure 7-14 takes a snapshot view of MPLS, and Figure 7-15 depicts the operation of MPLS.

Figure 7-13 depicts two data flows from Host X: one to Y and one to Z. Two LSPs are shown:

- LSR A is the ingress point into the MPLS network for data from Host X. When it receives packets from X, LSR A determines the FEC for each packet, deduces the LSP to use, and adds a label to the packet. LSR A then forwards the packet on the appropriate interface for the LSP.
- LSR B is an intermediate LSR in the MPLS network. It simply takes each labeled packet and uses the pairing {incoming interface, label value} to decide the pairing {outgoing interface, label value} with which

Figure 7-13
Two LSPs in an MPLS packet-switched network.

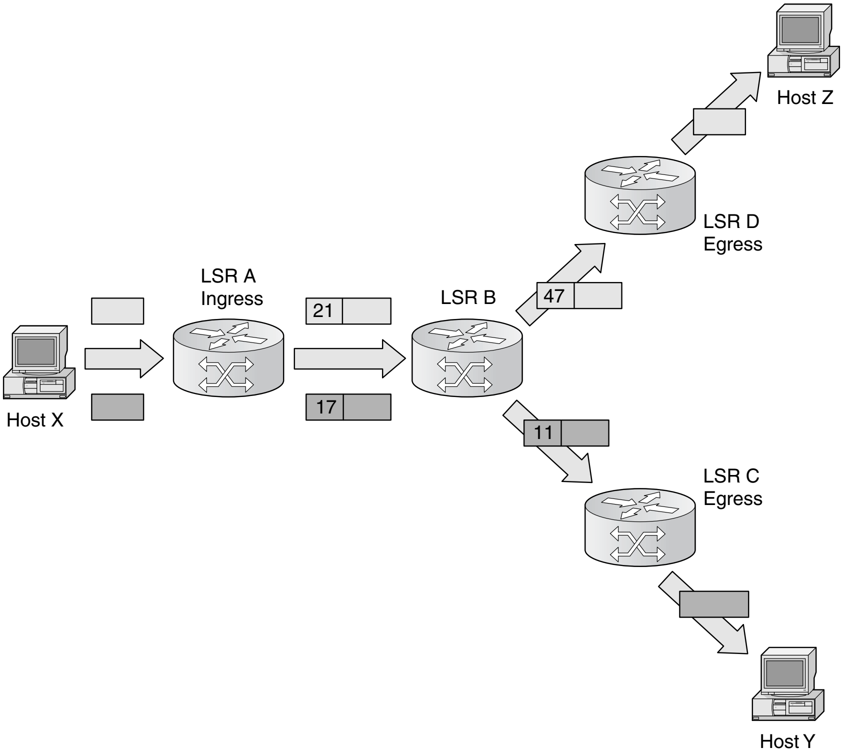
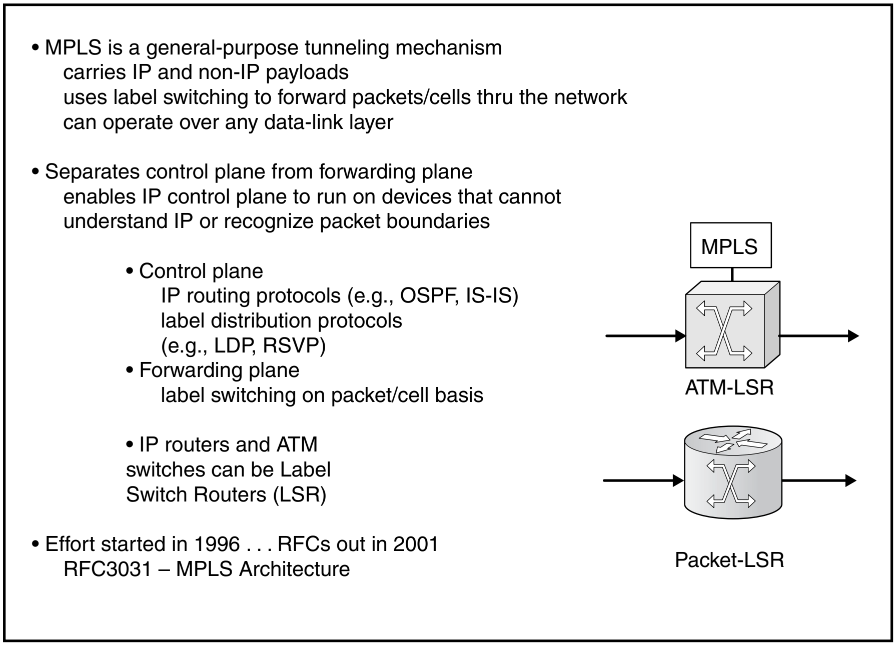


Figure 7-14
MPLS in a nutshell.
Source: Cisco.



- MPLS is a general-purpose tunneling mechanism
  carries IP and non-IP payloads
  uses label switching to forward packets/cells thru the network
  can operate over any data-link layer
- Separates control plane from forwarding plane
  enables IP control plane to run on devices that cannot understand IP or recognize packet boundaries
  - Control plane
    IP routing protocols (e.g., OSPF, IS-IS)
    label distribution protocols (e.g., LDP, RSVP)
  - Forwarding plane
    label switching on packet/cell basis
  - IP routers and ATM switches can be Label Switch Routers (LSR)
- Effort started in 1996 . . . RFCs out in 2001
  RFC3031 – MPLS Architecture

to forward the packet. This procedure can use a simple lookup table and can be performed in hardware, along with the swapping of the label value and forwarding of the packet. This enables MPLS networks to be built on existing label-switching hardware such as ATM and Frame Relay. This way of forwarding data packets can be potentially much faster than examining the full packet header to decide the next hop. In the example, each packet with the label value 21 will be dispatched out of the interface toward LSR D, bearing the label value 47. Packets with the label value 17 will be relabeled with the value 11 and sent toward LSR C.

- LSR C and LSR D act as egress LSRs from the MPLS network. These LSRs perform the same lookup as the intermediate LSRs, but the {outgoing interface, label value} pair marks the packet as exiting the LSP. The egress LSRs strip the labels from the packets and forward them using Layer 3 routing. So, if LSR A identifies all packets for Host Z with the upper LSP and labels them with the value 21, they will be successfully forwarded through the network. Note that the exact format of a label and how it is added to the packet depends on the Layer 2 link technology used in the MPLS network. For example, a label could correspond to an ATM Virtual Path Identifier/Virtual

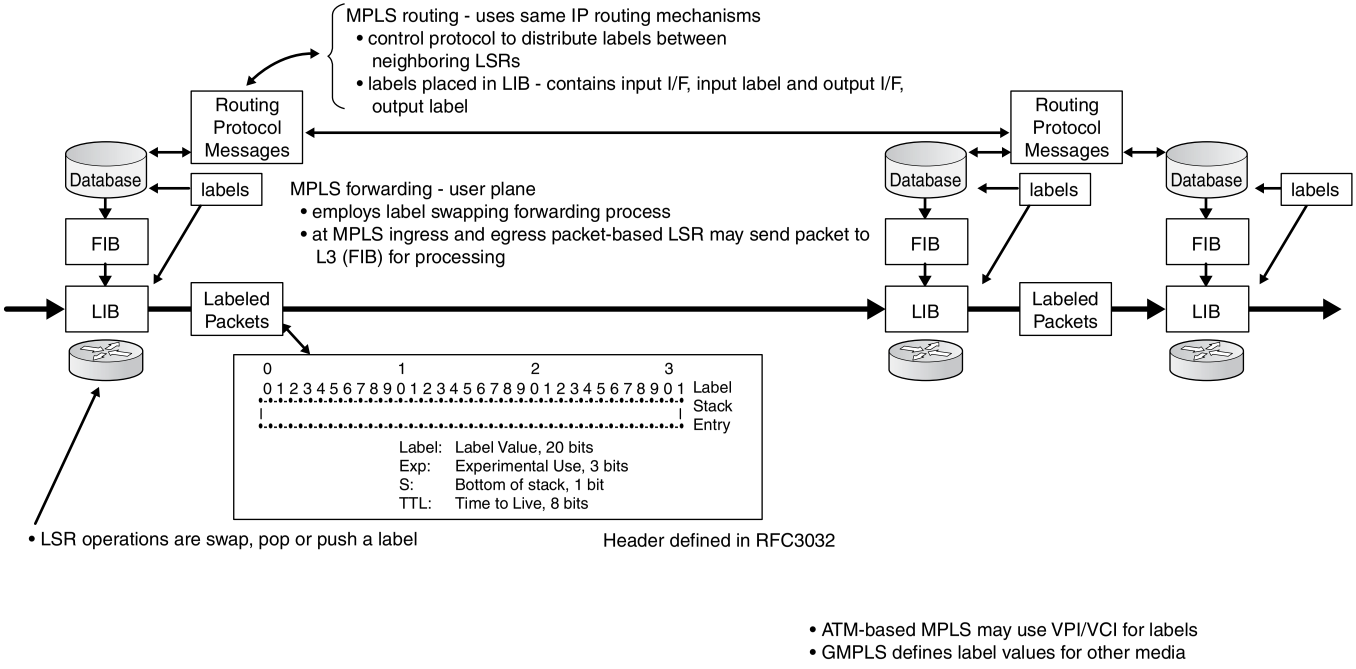


Figure 7-15
MPLS operation.

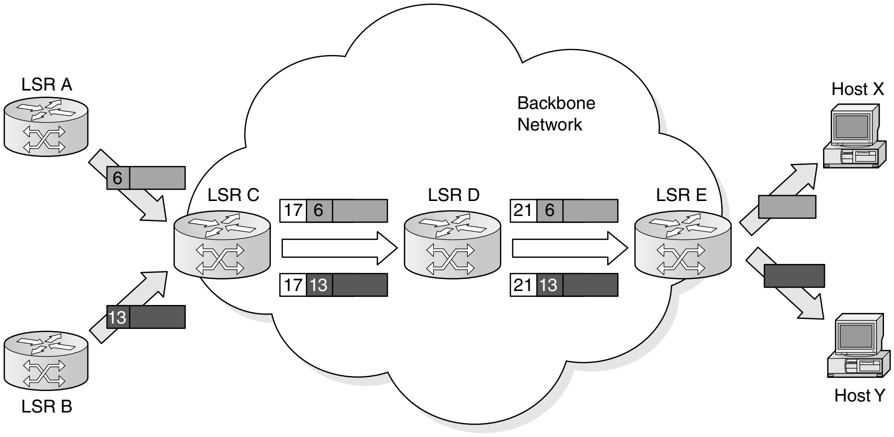
Channel Identifier (VPI/VCI) or a Frame Relay Data Link Connection Identifier (DLCI). For other Layer 2 types (such as Ethernet and PPP), the label is added to the data packet in an MPLS *shim* header, which is placed between the Layer 2 and Layer 3 headers. In a similar way, a label could correspond to a fiber, a DWDM wavelength, or a TDM time slot.

In order for LSPs to be used, the forwarding tables at each LSR must be populated with the mappings from {incoming interface, label value} to {outgoing interface, label value}. This process is called *LSP setup*, or label distribution. The MPLS architecture document does not mandate a single protocol for the distribution of labels between LSRs. The MPLS architecture document allows multiple different label LDPs for use in different scenarios. Alternatively, LSPs may be configured as *static* or *permanent* LSPs by programming the label mappings at each LSR on the path using some form of management such as Simple Network Management Protocol (SNMP) control of the Management Information Bases (MIBs).

A key to MPLS is that once the labels required for an LSP have been exchanged between the LSRs that support the LSP, intermediate LSRs transited by the LSP do not need to examine the content of the data packets flowing on the LSP. For this reason, LSPs are often considered to form tunnels across all or part of the backbone MPLS network. A tunnel carries opaque data between the tunnel ingress and egress LSRs. In Figure 7-13, both LSPs are acting as tunnels. LSR B forwards the packets based only on the label attached to each packet. It does not inspect the contents of the packets or the encapsulated IP header. Where LSPs are parallel, they can be routed, together, down a higher-level LSP tunnel between LSRs in the network. Labeled packets entering the higher-level LSP tunnel are given an additional label to see them through the network and retain their first-level labels to distinguish them when they emerge from the higher-level tunnel. The process of placing multiple labels on a packet is known as *label stacking*, which is shown in Figure 7-16. Label stacks provide a finer granularity of traffic classification between tunnel ingress and egress nodes than what is visible to the LSRs in the core of the network, which route data solely on the basis of the topmost label in the stack. This helps to reduce both the size of the forwarding tables that need to be maintained on the core LSRs and the complexity of managing data forwarding across the backbone.

In Figure 7-16, two LSPs between LSR A and LSR E, and between LSR B and LSR E, shown as dark grey labels along the top path and darker gray labels along the bottom path, are transparently tunneled across the back-

Figure 7-16 Label stacks across the backbone.



bone network in a single outer LSP between LSR C and LSR E. A label stack is arranged with the label for the outer tunnel at the top and the label for the inner LSP at the bottom. On the wire (or fiber), the topmost label is transmitted first and is the only label used for routing the packet until it is popped from the stack and the next highest label becomes the top label.

A major asset of MPLS is its traffic engineering. According to *draft-ietf-tewg-principles-00.txt*, " . . . Internet traffic engineering is defined as that aspect of Internet network engineering dealing with the issue of performance evaluation and performance optimization of operational IP networks." The objectives are to enhance performance, facilitate reliable network operations, and control and optimize the routing function. In other words, the goal is to forward packets over a path to the destination that offers the maximum efficiency and optimal performance. MPLS traffic engineering overrides shortest path routing. It discovers the path that meets requirements and installs the forwarding state along that path. Essentially, it provides connection-oriented service over a connectionless network. MPLS traffic-engineering components include routing protocol extensions, constraint-based path selection, signaling, and label switching. The advantages afforded by MPLS traffic engineering are better control over traffic aggregate paths, extensible to circuit-oriented network regimes, and polymorphic Link State Protocol (LSP) behavior. This is a brief description of how MPLS traffic engineering works:[17]

1. LSP advertises "unreserved capacity" and administrative attributes per link.
2. Constraints (required bandwidth and policy) are specified for a traffic-engineering trunk at the ingress mode.

3. To route a trunk, prune the unsuitable links from the topology.
4. Pick the shortest path on the remaining topology.
5. Use RSVP extensions to route the trunk and perform admission control.
6. MPLS handles the forwarding.

7.3.2 Optical Elements

In general, the point-to-point links between optical switches consist of bundles of optical fibers. An optical switch can choose to switch any possible subset of the optical bundle as a single unit—all the data traffic in such a unit on the incoming interface is switched to a corresponding type and size of unit on the outgoing interface. The switching variants used in optical networks are described in the following section. A key feature of all of them is that the optical switch switches large flows of data as a unit and does so based on quantities (wavelengths, time slots, and so on) inherent in the network medium rather than by examining data headers at the individual packet or application level.

Switching Entire Fibers A basic entity to switch is an entire fiber. All of the data that arrives on a single fiber is switched to be transmitted out of another fiber.

Lambda Switching Within a single fiber, the available bandwidth can be divided up by frequency into wavelengths (also known as *lambdas*). An optical switch could switch all of the data in wavelength A on the incoming fiber to wavelength B on the outgoing fiber. A restriction of some optical switches (such as MEMS) is that they are incapable of wavelength conversion, in which case B must be the same as A. Note that this is still different from switching the entire fiber because two different wavelengths on a single incoming fiber could be switched to retain their wavelength values, but exit the optical switch on two different outgoing fibers.

Waveband Switching Waveband switching is a generalization of lambda switching. If a fiber's bandwidth is divided by frequency, wavelengths may be grouped together and switched as a block. This has potential benefits in reducing the number of LSPs in place, which saves on signaling and switching hardware. The process of switching a waveband can be viewed as an LSP tunnel that switches each of the payload wavelength LSPs in the same way. Additionally, waveband switching may help to reduce the optical distortion that may be introduced by separating out and switching the individual lambdas.

Time Division Multiplexing (TDM) A fiber's bandwidth can also be divided by time slots. In this model, the optical signal is seen as a sequence of data frames, with N frames each of size S traveling every second (making up the total fiber bandwidth $N \times S$), and bandwidth is allocated for a particular data flow by reserving a portion of each frame. The basic frame sizes and the hierarchies by which a single frame can be divided into time slots are the subject of several standards, notably SONET and SDH.

Switching Quantities, Bandwidth, and Quality of Service (QoS) Switching quantities in optical networks are inextricably linked with bandwidth and QoS. Bandwidth is determined exactly by the type and size of the switching unit reserved (one fiber, one wavelength, one OC-48, and so on). QoS boils down to this bandwidth together with complete reliability and very low burstiness/jitter. In Chapter 8, we discuss how optical-switching quantities are interpreted and represented as GMPLS labels and how limitations on the types of switching that an optical switch is physically able to perform translate into constraints on the selection of GMPLS label values.

Out-of-Band Signaling Because all the data traffic passing through an optical switch can be switched without reference to individual data packets, there is no need for the data plane part of the optical switch to have any understanding of the protocol stacks (IP, Transmission Control Protocol [TCP], User Datagram Protocol [UDP], and so on) that are needed for the handling of control messages. In particular, it may not be necessary for optical switches to electronically terminate individual links.

Between a pair of core optical switches, there may be multiple data links. It could be both wasteful and confusing to establish a separate in-band signaling session within each link. It is more efficient to manage the links as a group using a single out-of-band signaling session. Typically, therefore, an optical switch completely separates its control plane from its data plane, and control connections to other switches go via a lower-performance, nonoptical network.

Performance Characteristics One great flexibility of MPLS networks is their capability to define a hierarchy of LSPs over which data is passed.

This enables more transitory, low-bandwidth LSPs, starting and finishing close to the network edge, to use preexisting high-bandwidth LSPs that span the core of a provider network. The requirement for high-bandwidth, long-distance LSPs, coupled with their long lifetime, forms a natural fit with optical technologies. Consequently, LSPs in optical networks are typically long lived and stable. On the other hand, optically switched LSPs can be slow to set up compared with electronically switched LSPs because micromirrors need time to be physically adjusted and to wait for the resulting movement vibrations to damp away. The stability and slow setup of optical MPLS networks have influenced the development of GMPLS.

Bidirectional LSPs Trunks through the core of an optical network are typically bidirectional. GMPLS is needed for them to set up bidirectional LSPs.

Management Issues By their nature, the mechanisms that optical switches use to switch traffic flows—such as micromirrors—do not notice if the traffic flow disappears altogether. Without additional hardware and software support, therefore, a break in the core of an optical network might not be detectable until the egress of that network, where the egress LSR tries to convert the signal back to packet form and finds it to be missing. This is known as the *loss of light (LoL)* problem. Fault localization in an optical network requires that the switches can be asked, under management control, to check for LoL on their incoming fibers. Conversely, where an LSP is being torn down gracefully under the control of a signaling protocol, the LoL condition propagates along the fiber much more quickly than the signaling messages that inform each switch of the situation. In this situation, any downstream OXC that is capable of LoL detection must be prevented from raising a false alarm about signal loss. Other management problems in an optical network include communicating information between switches about how optical fiber bundles are addressed at either end of a link and setting up backup links to kick in if a primary link fails.

Optical Network Client Interface Currently, optical networks are mostly confined to network cores with existing lower-cost networking technologies bridging the gap between core and edge networks. Therefore, an optical network needs to provide an interface to enable client network devices—such as routers—to dynamically request connections through it over the Optical User-to-Network Interface (O-UNI).

End Notes

1. For a thorough discussion of these kinds of signaling, see, for example, D. Minoli and G. Dobrowski, *Signaling Principles for Frame Relay and Cell Relay Services* (Boston: Artech House, 1994).
2. R. A. Barry, "Intelligent Optical Networking, Year 2001: A Practical Perspective," Next-Generation Networks Conference (NGN) 2001 Proceedings.
3. Others, however, state that these proponents prefer the simple diagrams of a few nodes (n) with a few links between them (say, $n+m$) because these diagrams are much simpler to deal with and to talk around, compared with the reality of the networks that are actually deployed (which indeed look more like the cascaded-ring diagrams that we show throughout this book). Also, because SONET/SDH is historically well suited and optimized for a ring configuration, a way to disentrench SONET is to deprecate the architecture it supports so well —namely, rings.
4. It is unclear how these numbers are derived. For example, it would take 500,000,000 simultaneous calls (1 billion people talking on the phone) to generate 30 Tbps at 64,000 bps DS0s.
5. C. Metz, Cisco Materials.
6. Extending the reference/experience horizon from 12 months in the past and 18 months into the future would help technology providers bring better and more revenue-generating products to the market.
7. NP-complete refers to the *complexity class* of *decision problems* for which answers can be checked for correctness, given a *certificate*, by an algorithm whose run time is polynomial in the size of the input (that is, it is *NP*) and no other NP problem is more than a polynomial factor. Informally, a problem is NP-complete if answers can be verified quickly and a quick algorithm to solve this problem can be used to solve all other NP problems.
8. Gee-Kung Chang, "Optical Label Switching," DARPA NGI PI Meeting, McLean, VA, October 4, 2000.
9. Often technology vendors make the pitch that the cost of bandwidth goes down using this or that technology. However, there is a procedural disconnect. In making those pitches, they are talking to the end user. Yet it is the carriers that are going to purchase the equipment these vendors manufacture. Therefore, the equipment vendors have to talk to

the carrier; they have to stimulate the interest of and show advantages and gains to the carriers.

10. O. Gerstel and R. Ramaswami, "Optical Layer Survivability: A Services Perspective," *IEEE Communications Magazine* 38, no. 3 (March 2000): 104–113.
11. X. Fang, R. Iraschko, and R. Sharma, "All-Optical Four-Fiber Bidirectional Line-Switched Ring," *IEEE Journal of Lightwave Technology* 17, no. 8 (August 1999): 1,302–1,308.
12. Adva Optical Networking White Paper, "Applications of Switch Matrices in Metro Optical Networking," May 15, 2001.
13. Neil Jerram and Adrian Farrel, "MPLS in Optical Networks—An Analysis of the Features of MPLS and Generalized MPLS and Their Application to Optical Networks, with Reference to the Link Management Protocol and Optical UNI," Data Connection Ltd. White Paper, Version 2, October 2001.
14. Goodness-of-fit tests indicate whether or not it is reasonable to assume that a random sample comes from a specific distribution. Statistical techniques often rely on observations that have come from a population that has a distribution of a specific form (for example, normal, lognormal, Poisson, and so on). There may be historical or theoretical reasons to assume that a sample comes from a particular population as well. Past data may have consistently fit a known distribution, for example, or a theory may predict that the underlying population should be of a specific form. Three common goodness-of-fit tests are the Chi-square test for continuous and discrete distributions, the Kolmogorov-Smirnov test for continuous distributions based on the empirical distribution function (EDF), and the Anderson-Darling test for continuous distributions (http://www.itl.nist.gov/div898/handbook/prc/section2/prc21.htm). A more extensive treatment of goodness-of-fit techniques is presented in R. B. D'Agostino and M. A. Stephens, *Goodness-of-Fit Techniques* (New York: Marcel Dekker, Inc., 1986). Also see C.R. Hicks, *Fundamental Concepts in the Design of Experiments* (New York: Holt, Rhinehart, and Winston, 1973); R.G. Miller Jr., *Simultaneous Statistical Inference* (New York: Springer-Verlag, 1981); Neter, Wasserman, and Whitmore, *Applied Statistics* 4th ed., (Boston: Allyn and Bacon, 1993); and H. Scheffe, *The Analysis of Variance* (New York: John Wiley, 1959).
15. Portions of this section are based directly on Neil Jerram and Adrian Farrel's "MPLS in Optical Networks—An Analysis of the Features of

MPLS and Generalized MPLS and Their Application to Optical Networks, with Reference to the Link Management Protocol and Optical UNI," Data Connection Ltd. White Paper, Version 2, October 2001.

16. Do not confuse this with forward error correction (FEC).
17. C. Metz, "A Survey of Advanced Internet Protocols," Cisco Educational Materials, 2001.

CHAPTER 8

GMPLS in Optical Networks

One of the most exciting recent developments in optical networks has been the attempt to standardize a generalized version of the multiprotocol label switching (MPLS) architecture so that it can be used to signal paths through an optical network.

In a classical MPLS network, incoming packets are assigned a label by a label edge router (LER). Packets are forwarded along a label-switched path (LSP) where each label-switched router (LSR) makes forwarding decisions based solely on the contents of the label. At each hop, the LSR strips off the existing label and applies a new one that tells the next hop how to forward the packet. RFC 3031 "MPLS Architecture" has a detailed explanation of this process.

However, classical MPLS labels are carried in the header of each cell or packet that flows through the MPLS network. Optical switches, on the other hand, forward large streams of data without decomposing them into packets or cells so the header-based labels of classical MPLS cannot tell an optical switch how to switch that data flow. Hence, the need for MPLS to be generalized, as the Internet Engineering Task Force (IETF) specification of Generalized MPLS (GMPLS) explains, is as follows.[1]

The original architecture has recently been extended to include LSRs whose forwarding plane recognizes neither packet nor cell boundaries, and therefore cannot forward data based on the information carried in either packet or cell headers. Specifically, such LSRs include devices where the forwarding decision is based on time slots, wavelengths, or physical ports.

Given the previous information, LSRs (or interfaces on LSRs) can be subdivided into the following groups:

- **Interfaces that recognize packet/cell boundaries and can forward data based on the content of the packet/cell header** Examples of these include interfaces on routers that forward data based on the content of the shim header and that are interfaces on Asynchronous Transfer Mode (ATM) LSRs that forward data based on the ATM Virtual Path Identifier/Virtual Channel Identifier (VPI/VCI). Such interfaces are referred to as packet-switch capable (PSC).
- **Interfaces that forward data based on the data's time slot in a repeating cycle** An example of such an interface is an interface on a Synchronous Optical Network (SONET) crossconnect. Such interfaces are referred to as time-division multiplex capable (TDMC).
- **Interfaces that forward data based on the wavelength on which the data is received** An example of such an interface is an

interface on an optical crossconnect (OXC) that can operate at the level of an individual wavelength. Such interfaces are referred to as lambda-switch capable (LSC).

- **Interfaces that forward data based on a position of the data in real-world physical spaces** An example of such an interface is an interface on an OXC that can operate at the level of single or multiple fibers. Such interfaces are referred to as fiber-switch capable (FSC).

Using the concept of a nested LSP enables the system to scale by building a forwarding hierarchy. At the top of this hierarchy are FSC interfaces, followed by LSC interfaces, followed by TDMC interfaces, followed by PSC interfaces. This way an LSP that starts and ends on a PSC interface can be nested (together with other LSPs) into an LSP that starts and ends on a TDM interface. This LSP, in turn, can be nested (together with other LSPs) into an LSP that starts and ends on an LSC interface, which can be nested (together with other LSPs) into an LSP that starts and ends on an FSC interface.

The establishment of LSPs that span only the first class of interfaces is defined in the Label Distribution Protocol (LDP), constraint-based routing LDP (CR-LDP), and Resource Reservation Protocol Traffic Engineering (RSVP-TE). The GMPLS specification presents a functional description of the extensions needed to generalize the MPLS control plane to support each of the four classes of interfaces. Only signaling-protocol-independent formats and definitions are provided in this document. Protocol-specific formats are defined in the documents GMPLS-RSVP and GMPLS-LDP. Technology-specific details are specified in technology-specific documents, such as GMPLS-SONET.

The sections that follow, which are contributed by Data Connection Ltd., explain how the concepts of classical MPLS are generalized enough to be applicable to optical networks and describe key aspects of GMPLS technology.[2]

8.1 Labels in Optical Networks

A basic requirement in MPLS is that the two LSRs at either end of a link must agree on how they will mutually identify a traffic flow. For this purpose, they use a label that is assigned by one of the LSRs and distributed to the other LSR using signaling protocol messages.

8.1.1 Nongeneralized Labels

In nongeneralized MPLS, a label is a number (up to 32 bits) that is written into the protocol header fields of data packets traveling on the link. Once a pair of LSRs have agreed on this number, they also agree that the corresponding data flow consists of all data packets with this number in the appropriate protocol header field and will switch all the packets in this flow in the same way.

Note that with this kind of label, the agreed label value does not necessarily imply a relationship to bandwidth allocation or quality of service (QoS) for the corresponding data flow. The label value might not imply anything about how frequently packets with that value might arrive or what bandwidth is available.

Traffic-engineering LDPs (such as RSVP and CR-LDP) facilitate QoS and bandwidth negotiation as part of the label exchange. Other protocols (such as LDP) simply exchange labels.

8.1.2 Generalized Labels

The premise of GMPLS is that the idea of a label can be generalized to be anything that is sufficient to identify a traffic flow. For example, in an optical fiber whose bandwidth is divided into wavelengths, the whole of one wavelength could be allocated to a requested flow—the LSRs at either end of the fiber simply have to agree on which frequency to use. Unlike nongeneralized labels, the data inside the requested flow does not need to be marked with a label value; instead, the label value is implicit in the fact that the data is being transported within the agreed frequency band. On the other hand, some representation of the label value is needed in the signaling protocol so that control messages between the LSRs can agree on the value to use.

GMPLS extends the representation of a label from a single 32-bit number to an arbitrary-length byte array and introduces the Generalized Label object (in RSVP) and Generalized Label Type, Length, and Value (TLV) (in CR-LDP) to carry both the label and related information. The following subsections describe how the switching quantities used in optical networks are represented as GMPLS labels.

Whole Fiber Labels A link between LSRs may consist of a bundle of optical fibers. LSRs may choose to allocate a whole fiber to a data flow so they need to agree on which fiber (within the bundle) to use. In this case, the label value is the number of the selected fiber within the bundle. The interpretation of the fiber/port numbers is a local matter for the LSRs on the link. If the two LSRs use different numbering schemes, Link Management Protocol (LMP)[3] provides a mechanism for LSRs to exchange and correlate numbering information—see Section 8.6.

Wavelength Labels Where the bandwidth of an optical fiber is subdivided by wavelength division multiplexing (WDM), an optical LSR may choose to allocate a single wavelength—or lambda—to a requested data flow. In this case, the label value is the wavelength of the selected lambda.

Waveband Labels If consecutive wavelengths are grouped together into a waveband, in order to be switched in the same way, the label is a waveband ID and a pair of numbers (channel identifiers) that indicates the lower and upper wavelengths of the selected waveband.

Time Slot Labels When the bandwidth of an optical fiber is subdivided into time slots by TDM, an optical switch may satisfy a particular data flow request by allocating one or more time slots to that flow. In general, therefore, a TDM label value must be sufficient to specify the allocated time slot(s). The exact details of TDM label representation depends upon the TDM hierarchy in use—for example, SONET or Synchronous Digital Hierarchy (SDH).

SONET/SDH Labels A SONET/SDH label is represented as a sequence of five terms, known as S, U, K, L, and M, which select branches of the SONET/SDH TDM hierarchy at increasingly fine levels of detail.

Bandwidth Allocations For all the types of GMPLS labels described here, the label value directly implies the bandwidth that is available for the corresponding data flow. For example, if a label denotes a single SONET VT-6 time slot, the available bandwidth is, inextricably, the bandwidth of a VT-6 time slot. The same holds true for other TDM labels and lambda, waveband, or fiber labels. This is quite different from the case for non-generalized labels and is a fundamental reflection of the nature of optical networks.

8.1.3 Requesting Generalized Labels

The basic MPLS protocol for agreeing a label value across a link is unchanged for optical networks:

- The upstream LSR sends a request to the downstream LSR (a Path message in RSVP and a Label Request message in CR-LDP). The request contains enough information about the requested bandwidth and QoS for the downstream LSR to make a sensible label choice.
- The downstream LSR receives the request and allocates a label value that meets the requirements specified by the request.
- The downstream LSR sends a response to the upstream LSR (the Resv message in RSVP and label mapping in CR-LDP) that communicates the selected label value.

GMPLS generalizes the Request Setup message for two reasons: to distinguish it from a nongeneralized request setup and to enable it to carry additional parameters that specify the request in more detail. In RSVP, this is done by using a Generalized Label Request object instead of a LABEL_REQUEST object in the Path message and CR-LDP by adding a Generalized Label Request TLV to the Label Request message.

Some of the information that the downstream LSR needs in order to allocate a suitable label value is implied by the context. At the most basic level, both LSRs know that the label must be a generalized one, rather than, say, a nongeneralized ATM label, because they know that the link to which the request applies is a generalized optical link. Therefore, this information is not explicit in the request message.

However, because an optical link may consist of a bundle of fibers and the switches may support more than one kind of multiplexing on those fibers, it is necessary for the upstream LSR to specify the LSP encoding type that it wants for the data flow being set up. This encoding type then determines whether the agreed label will be time slot or wavelength based and determines the kind of label.

So the Generalized Label Request specified by GMPLS carries an LSP encoding-type field. The currently supported values for this field that are relevant to optical networks are

- American National Standards Institute (ANSI) Plesiochronous Digital Hierarchy (PDH)
- European Telecommunications Standards Institute (ETSI) PDH
- SDH

- SONET
- Digital wrapper
- Lambda
- Fiber

Because some links may advertise the capability to support more than one type of switching capability (through the Interior Gateway Protocol [IGP]), the Generalized Label Request object/TLV contains a field that indicates the switching mode to be applied to the particular LSP. This enables, for example, a switch to switch whole fibers, wavebands, or individual lambdas. The choice of how to switch for any particular LSP is made when the LSP is set up. This increases the flexibility of how the network resources can be used. For fiber- and wavelength-based labels, nothing more is needed.

When requesting SONET and SDH labels, it may be necessary to request that the total bandwidth for the LSP be split across multiple time slots. Therefore, when the LSP encoding type is SONET or SDH, the Generalized Label Request carries additional fields that specify how many time slots should be combined to meet the request (the Requested Number of Components [RNC] field) and how those time slots should be concatenated, including whether they are required to be contiguous (the Requested Grouping Type [RGT] field).

8.2 Constraining Label Choice

As described in the previous section, the choice of label for each link is normally dictated by the downstream node on that link. In nongeneralized MPLS, where labels are simply arbitrary numbers (such as ATM VPI/VCIs), this is fine. In GMPLS, where labels are directly related to network resources, this can lead to conflicts during LSP setup.

For example, an optical switch based on micromirrors may be able to switch a received wavelength from an incoming port to an outgoing port, but may not be able to modify the wavelength.

Figure 8-1 shows two LSPs in an optical network where the switches (OXCs) are incapable of wavelength conversion. The solid line LSP runs from optical switch A through B, D, and E to F. The dashed line LSP runs from C through D and E to G. The different line types also represent different wavelengths in use for the LSPs, where the solid line represents the blue wavelength and the dashed line represents the red wavelength. There

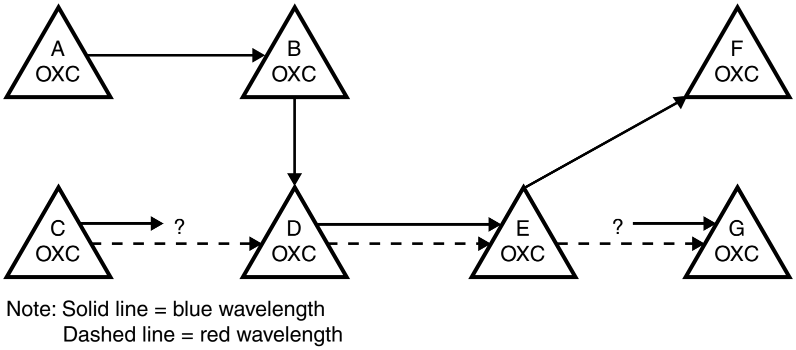


Figure 8-1 Conflict of labels at optical switches.

is no conflict on the link between D and E because the two LSPs use different wavelengths.

However, when LSR C needs to set up a new LSP through D and E to G, it must pick a new wavelength. If the choice is left to LSR G (which would be the normal way of processing in MPLS), G might choose the solid wavelength (blue). But the solid wavelength (blue) is already in use between D and E so it cannot be used. If the choice is given to LSR C, a similar problem may occur. There is therefore a need to allow all OXCs along the path to constrain and/or influence the choice of labels to ensure that appropriate labels are chosen.

8.2.1 Label Set

GMPLS introduces the concept of a label set. An upstream LSR includes a label set on its signaling request to restrict the downstream LSR's choice of label for the link between them. The downstream LSR must select a label from within the label set or else it must fail the LSP setup.

This is useful in the optical domain when, for example,

- An LSR is incapable of converting between wavelengths, as shown in the previous example.
- An LSR can only generate and receive a subset of the wavelengths that can be switched by neighboring LSRs.
- It is desirable for an LSR to limit the amount of wavelength conversion that it has to perform in order to reduce the distortion on the optical signals.

The label set is constructed by including and/or excluding an arbitrary number of label lists and/or ranges. If no labels are explicitly included, the set consists of all of the valid labels not explicitly excluded. If no label set is present, the downstream LSR is not constrained in its choice of label.

As the label set is propagated on the Path message, each LSR may generate a new outgoing label set based on its hardware capabilities and possibly on the incoming label set. For example, an LSR that is incapable of wavelength conversion generates an outgoing label set by forwarding the incoming label set minus any labels that it cannot generate or receive. In contrast, an LSR that is capable of converting all wavelengths in an incoming label set may choose to remove the label set when sending the signaling request downstream to indicate that it will accept any label for use on its downstream link.

Consider the example in Figure 8-1. Suppose LSR C can only generate wavelengths from a range R. Therefore, it signals a label set of "anything in range R except red." LSR D modifies this and signals "anything in range R except red or blue." LSR E forwards this unchanged, and LSR G can select a color that will be acceptable to all of the LSRs in the path, if one is available.

Note that the color is not guaranteed to be acceptable. If LSR G chooses, say, green, another LSP B-D-C using green may have been set up while LSR D was waiting for the signaling response from LSR E. This risk can be reduced by also using suggested labels—see the section titled "Suggested Label."

8.2.2 Explicit Label Control

GMPLS also introduces explicit label control. This enhances the MPLS concept of an explicit route by enabling the ingress LSR to specify the label(s) to use on one, some, or all of the explicitly routed links for the forward and/or reverse path.

This is useful, for example, when the ingress LSR wants to insist that the wavelength used is the same along the whole LSP. This might be desirable in order to avoid distortion of the optical signal.

It may also be useful in traffic engineering where the path computation engine has knowledge of the labels in use in the network and the switching capabilities of the LSRs. In this case, the path can be computed to include the specific labels to be used at each hop.

Explicit labels are specified by the ingress LSR as part of the explicit route. At each LSR along the path, any explicit label that is specified in the explicit route for the next hop is removed and converted into a Label Set object, containing a single label, for the next hop. The LSR that receives this label is then required to use it for that hop and must fail the setup if that label is not locally available.

8.2.3 Egress Label Control

When a network administrator initiates the setup on an LSP, he or she is free to specify more or less strictly the path that the LSP should follow and, optionally, as described in the previous subsection, the label values to use on the links that the LSP traverses.

Sometimes the network administrator may have additional information about the routing of data traffic as it emerges from the far end of the LSP. For example, he or she may know that only voice telephony traffic will be injected into this LSP, and that upon exit from the LSP, all of this traffic should be routed via a telephony gateway with a well-known address. In such cases, it makes sense to avoid the data packet examination and routing calculation that the egress LSP would otherwise have to perform by signaling the well-known gateway address along the LSP.

The mechanism to achieve this is called *egress label control*. The administrator adds additional label objects to the explicit route after the last-hop object. The encoded labels do not have to conform to any standard MPLS label format, but can be anything that makes sense to the egress LSR. When the egress LSR receives the LSP setup message, it notices the additional label objects at the end of the explicit route and interprets them in whatever way it finds useful.

8.3 Out-of-Band Signaling

The nongeneralized MPLS signaling protocols presume that the data traffic in an LSP will follow the same path as the signaling messages. In optical networks, however, the bandwidth granularity of channels in optical links is high, and it would be wasteful to use an entire bandwidth slice (time slot or wavelength) as a signaling channel. Therefore, there are strong reasons for the signaling to follow an *out-of-band* path instead of via a control channel that is physically distinct from the data channel. It also simplifies the technology that an optical switch needs to implement in its data plane, if the data plane does not need to understand the protocols upon which signaling messages are based.

Some solutions use a low-bandwidth link (such as Ethernet) running in parallel to the data channel. Alternatively, it may be that the need to provision a dedicated signaling channel can be avoided by routing via an existing Internet Protocol (IP) cloud.

Out-of-band signaling raises three key issues for the application of GMPLS to optical networks:

- The routing that an optical LSR performs during LSP setup must be extended to calculate suitable distinct next-hop IP addresses and outgoing interfaces for the data and signaling.
- Signaling messages may need to be encapsulated to ensure that they arrive successfully at the intended next-hop LSR.
- Because signaling messages are no longer in band, they need a way of indicating the data interface to which they refer.

The following subsections explain these issues further and how they are resolved.

8.3.1 Extending Routing Calculation

When an LSR on the data and signaling paths attempts to route the LSP setup, it must calculate two outgoing routes to the next hop: one via the data path and one via the signaling path. The data path must be evaluated first and a signaling path must then be found that reaches the next hop on the data path.

The topologies for the signaling and data networks are obviously different so the routing decision is complicated. Work is ongoing on optical extensions to Open Shortest Path First (OSPF) that will permit the distribution of topology data for both signaling and data networks. Once this data has been distributed, each individual LSR has the information that it needs to calculate the required routes for out-of-band signaled GMPLS.

8.3.2 Signaling Message Encapsulation

When running RSVP signaling, each signaling message is addressed to the egress (destination) LSR, never to the immediate next hop. Routers on the path intercept such messages by noticing that the Router Alert flag is set. In out-of-band signaling, there are two problems with this approach:

- The IP cloud spots a neat short route to the egress that entirely misses the next hop on the data path.
- A transit node in the IP cloud is RSVP capable, intercepts the signaling message, and attempts to process it.

There are two possible approaches to ensure that the signaling message arrives at the next hop on the data path and that the transit nodes in the signaling path do not interfere with it:

- The IP packet carrying the signaling message can be addressed to the LSR that is the next hop on the data path, rather than to the LSP egress, and can be sent without setting the Router Alert flag in the IP header. This is done for a couple of reasons:
  - Addressing to the next-hop LSR ensures that the packet will reach the correct LSR, but requires that the receiving LSR relaxes the rule that the IP packet should be addressed to the egress LSR. (In practice, many LSR implementations may not police this last rule anyway.)
  - Not setting the Router Alert flag should mean that the packet is forwarded without being intercepted by any router except the one to which it is addressed. However, it is likely that some IP stack implementations will not correctly check the Router Alert flag. When traveling through an arbitrary IP cloud, this is not a reliable solution.
- The IP packet can be sent *double encapsulated* with an extra IP header. The outer header is addressed to the next hop on the data path and shows the payload protocol as IP encapsulated, while the inner header is the normal header for an RSVP packet.

The second of these approaches is the more robust. It requires some additional straightforward function at the receiving LSR to recognize that it is the target of the IP packet, strip the outer header, and then treat the inner header as normal, passing it up to the signaling stack.

Note that these issues primarily concern RSVP signaling. CR-LDP transfers signaling data over Transmission Control Protocol/Internet Protocol (TCP/IP) sessions between signaling peers and therefore does not suffer from the problems described previously.

8.3.3 Data Interface Identification

An LSR can expect to receive out-of-band signaling messages over a different interface from the one that is going to carry data. This raises two concerns.

First, on receipt of a Path message, how does the LSR know which data interface is being signaled? There are several options:

- Instead of a network (IPv4 or IPv6) address, an explicit route may specify an unnumbered link object. The unnumbered link ID in this object is sufficient to identify the data interface to which the signaling message refers.
- An explicit route may specify the label for the hop. It is conceivable that the label value could also encode the interface index to which the label applies. For example, the label could carry the port ID in the top 16 bits and the lambda ID in the lower 16 bits.
- The data interface may be communicated via some new signaling protocol object. Second, the LSR needs to relax its explicit route processing rules so that it is acceptable for the top hop to refer to the data path and not the signaling path.

8.3.4 Standardization Status

The latest GMPLS drafts address some of the issues of out-of-band signaling:

- Data interface identification—this is supported by the interface ID TLV, which can carry any combination of the following kinds of addresses:
  - An IPv4 address.
  - An IPv6 address.
  - An IP address together with an interface identifier.
  - An IP address together with an upstream component link identifier.
  - An IP address together with a downstream component link identifier.
  - In Generalized RSVP-TE, this TLV is part of the IF_ID RSVP_HOP object, which replaces and extends the previously defined RSVP_HOP object.
  - In Generalized CR-LDP, this TLV forms the body of the new interface TLV.
- Control plane link and node failure.
- Generalized RSVP-TE uses standard RSVP refresh mechanisms to synchronize state following a control plane link or node failure, but it also does the following:
  - Defines a way to negotiate the maximum control plane recovery period

  - Reuses the suggested label object to enable an upstream node to tell the recovered downstream node about the label values that it allocated before failure
- Generalized CR-LDP uses the mechanisms already described in *draft-ietf-mpls-ldp-ft*[4] to synchronize signaling state following a link failure.

This just leaves the issue of how to ensure the delivery of RSVP-TE signaling messages to the next hop. (There is no corresponding issue for CR-LDP because CR-LDP uses TCP to send signaling messages.) Nothing is yet standardized for this question, but most current reports seem to favor the encapsulation approach, either IP in IP, as described previously, or Generic Routing Encapsulation (GRE), as described in RFC 2784.

8.4 Reducing Signaling Latency and Overhead

The performance characteristics of optical networks are often quite different from the electronically switched packet networks for which MPLS was originally developed:

- The time needed to set up an optical LSP may be longer than for a corresponding packet-switched LSP because of the mechanics of optical switch hardware.
- On the other hand, optical networks are particularly stable—once an LSP through an optical network has been established, it is likely to stay established for a long time.
- If an optical link fails, the discrepancy between the speed of the data path and the speed of signaling messages means that the standard MPLS error messages are not efficient tools for reporting and recovering from the failure. The following subsections explore how GMPLS has evolved to take these performance characteristics into account and work well with them.

8.4.1 Switch Programming Latency

The normal procedure for MPLS has the switch programmed when the signaling response is received. That is, the signaling request progresses through the network hop by hop from ingress to egress and then the signaling response travels from egress to ingress, causing the switch to be programmed as it goes. When the response reaches the ingress, the whole LSP is programmed and data may start to flow immediately.

Part of the reason for doing this is that labels are allocated by the downstream node on each hop, easing contention issues. Hence, on each hop, the label is not known by that hop's upstream LSR until the response is received.

Optical switches may be relatively slow to program. Although the time to select and adjust the switching components may be quite fast, the time taken for the components to settle down after programming can be much larger—measured in milliseconds. For example, a micromirror can be programmed quickly, but the mirror may take tens of milliseconds to stabilize and stop vibrating after it has been adjusted. It is not safe for an LSR to send a signaling response to its upstream neighbor while the mirror is still vibrating, as the ingress could then send data prematurely and the data would be lost or incorrectly switched.

Therefore, the time taken to establish an LSP that traverses n optical LSRs is $2\times$(end-to-end signaling time)$+n\times$(switch programming and settling time).

The combination of optical switches and conventional MPLS causes considerable latency in LSP setup.

Suggested Label To reduce the latency of LSP setup, GMPLS introduces the Suggested Label concept. Each LSR selects a label that it believes will be suitable for use on the link between itself and its downstream partner. It signals this label on the forward signaling path and immediately starts to program its own switch on the assumption that this label is the one that will be agreed on.

When the signaling response arrives back at the LSR, the message carries a label. If this label confirms the choice suggested on the request, nothing further needs to be done because the switch is already programmed. As long as the switch programming has settled down, the signaling response can be forwarded upstream immediately. If the label is different from the one suggested on the signaling request, the switch must be reprogrammed, but nothing is lost compared with the base case where no label was suggested.

In Figure 8-2, a signaling request is received at an LSR. For the sake of the example, the signaling protocol is RSVP-TE so the message is a Path.

The Path message carries a Suggested Label that indicates the label (λa) that the upstream LSR would likely use on the segment of the LSP that links the two LSRs. The receiving LSR processes the path as follows:

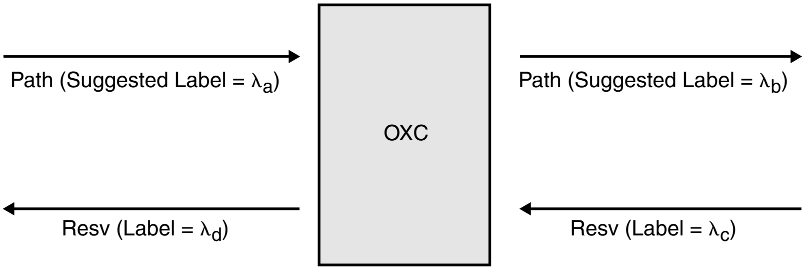


Figure 8-2
Label suggestion.

- It selects a label for use on the upstream link (λd). This should be the suggested label (λd=λa) if possible, but it can be any other label.
- It selects a preferred label for use on the downstream link (λb). Depending on the properties of the switch, this may be the same value as selected for the upstream interface (λd=λb), for example, if the switch is not capable of wavelength modification.
- It sends a command to the switch fabric to begin programming the switch (λd×λb).
- It sends a Path message downstream containing the label that it would likely use on the downstream segment (λb).
- At some point, it receives a Resv from the downstream LSR. This indicates the actual label to use on the downstream segment (λc).
- If the actual label is different from the suggested label (λc≠λb), the switch must
  - Possibly pick a new value for the upstream label (λd). This is necessary if the switch is not capable of wavelength conversion; it would therefore choose the same label as provided from downstream (λd=λc≠λa).
  - Send a command to the switch fabric to program the crossconnect (λd×λc).
  - Wait for the switch to be fully programmed and stable.
  - Send a command to the switch fabric to deprogram the speculative crossconnect (λd×λb) as a background event.
- If the label is the label suggested (λc=λb), then no further programming is required, but the LSR must make sure that the programming request made earlier has completed satisfactorily.
- Finally, the LSR sends a Resv upstream indicating the actual label to use on the link (λd).

If everything has worked well and wavelength conversion is not possible, the suggested label from the original message is used on all messages (λa=λb=λc=λd). If there are any problems (for example, an actual label is unacceptable), the LSP is torn using the normal processes for the signaling protocol and the switch is deprogrammed.

8.4.2 Soft-State Overhead

Optical MPLS networks are typically stable and the LSPs are long lived and relatively unchanging. This stability puts extra focus on the overhead of running RSVP, which requires regular refreshes between each network node to keep an LSP alive. The overhead of running this soft-state mechanism is more obvious in stable networks because the benefit of automatically cleaning up expired state is less frequently needed. The following sections outline some approaches to reducing this overhead.[5] All the approaches are general to RSVP-TE, but have particular applicability to optical networks because of their stability.

Note that soft-state overhead is not an issue for LDP and CR-LDP because they are hard-state protocols. The other side of the coin for LDP and CR-LDP is that they require special measures to preserve LSPs when a signaling session is temporarily broken.[4]

Message IDs The soft-state nature of RSVP allows the loss of protocol messages to be handled by simply waiting for the next refresh interval, or in the case of teardown messages, waiting for the state to time out. A consequence of this is that there is a tension between choosing a short refresh interval for robustness against message loss and the increasing overhead of running RSVP that this entails.

Message IDs provide an alternative approach to improving the robustness of message transmission that does not significantly increase the overhead of running RSVP. Each message includes a message ID, chosen by the sender, that uniquely identifies the message. Received message IDs are acknowledged by the receiver and are either piggybacked on other RSVP messages or sent in a lightweight ACK message if no other RSVP messages are available. The sender can then employ a retry timer for sent messages, resending messages until they're acknowledged by the recipient. As message loss is rare and ACKing message IDs is quick, message retries can be kept to a minimum, even with a relatively short interval for the retry timer. This mechanism gives the benefit of much improved reliability of message transmission, without significantly increasing the overhead of running RSVP.

Figure 8-3 illustrates the use of message IDs and ACKs.

Reduced Refresh Processing An important consequence of the message ID mechanism is that a given message ID can be used as a handle to the contents of the message:

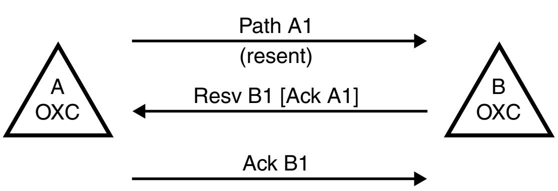
- If a Path or Resv message is being sent to refresh state in a neighboring node, then the same message ID is used again.
- When an LSR receives a Path or Resv with the same message ID as before, it indicates that the message is a simple refresh message, and the LSR can avoid processing the remaining message contents.

As the payload of RSVP messages increases, it becomes increasingly important to avoid the full processing of refresh messages.

Message IDs are 32-bit quantities. They are allocated by the sender in ascending order from some starting point (except in the case of refresh messages, as noted previously). This enables the receiver to detect and ignore messages that have arrived out of sequence. This mechanism also copes with nodes failing and potentially resetting the sequence numbers by including a randomly chosen *epoch* value on each message. If this value changes, peer nodes can interpret the message IDs appropriately.

Refresh Reduction Message IDs provide a simple way to reduce the processing load on the receiver by identifying refresh messages, but this still requires the sender to generate and send a complete RSVP message. Summary refresh (Srefresh) messages build on message IDs and provide a way to reduce the load on the sender and network as well. The sender batches up previously sent message IDs for state that it wants to refresh in a neighboring node, which could be Path state, Resv state, or a mixture

Figure 8-3
Message IDs in use.



of the two. The message IDs are then sent as a simple list in a dedicated Srefresh message.

On receipt of a Srefresh message, the receiver can quickly match the enclosed message IDs to its installed state, performing the normal keep-alive action for each state that is matched:

- Matched message IDs are not acknowledged.
- Unmatched message IDs are returned to the sender as unacknowledged in a NACK object. As with message ID ACKs, NACKs can be piggybacked on other RSVP messages or sent in an ACK message if no suitable RSVP message is available.
- When the Srefresh sender receives a NACK object, it matches it to the local state that generated that message ID and sends a full refresh message out immediately.

Srefresh provides a significant reduction in the overhead of running RSVP in stable networks, reducing the processing load on both sender and receiver and reducing the bandwidth required to transmit regular refresh messages. By replacing multiple Path and Resv messages with a single Srefresh message, the bandwidth savings can be large.

Although the reduction in the use of control channel bandwidth may not be important in optical networks where the control channel is often massively overprovisioned, the reduction in processing can be significant, especially in large networks with many LSPs.

Srefresh provides significant benefits for regular refresh messages, but cannot be applied to trigger messages, where the full message contents are required. Message bundling provides some performance enhancement in these situations by collecting the RSVP contents of several messages behind a single IP header. This saves some transmission bandwidth in IP (and possibly Layer 2) headers at the cost of some burstiness in transmission and added latency as messages are collated. It can provide benefits in failover scenarios where large numbers of teardown or setup messages are passing between neighbors.

8.4.3 Efficient Fault Handling

Recovery from failed network resources is an increasingly important aspect of network design. Ideally, the user should be protected from knowledge of a fault in the network. The user should not see any disturbance to his or her data traffic and should not have to take any remedial action.

MPLS offers many options for protecting LSPs from network problems.[6] In optical networks, these procedures utilize some specific extensions to MPLS, which are described in the following section.

Notify Messages Protection switching (as shown in Figure 8-4) is a scheme where a network failure is recovered by directing data from an interrupted primary LSP onto a typically presignaled backup LSP that uses a disjoint path.

In Figure 8-4, black and gray lines represent primary LSPs through the network. When the shared link in the middle of the network fails, data is redirected to flow on the backup LSPs (represented by the dashed lines). In order for this to work, an upstream node—the repair point—must become aware of the error. The repair point is usually the ingress/initiator, and in nongeneralized MPLS, the notification method is the PathErr message, which flows upstream following the path of the LSP. This message flow is effective, but is slow because of the following reasons:

- The path selected for the LSP may not be the shortest point between the detection point and the initiator (because of the traffic-engineering routing of the LSP), but the PathErr still follows this path delaying the notification of the error.

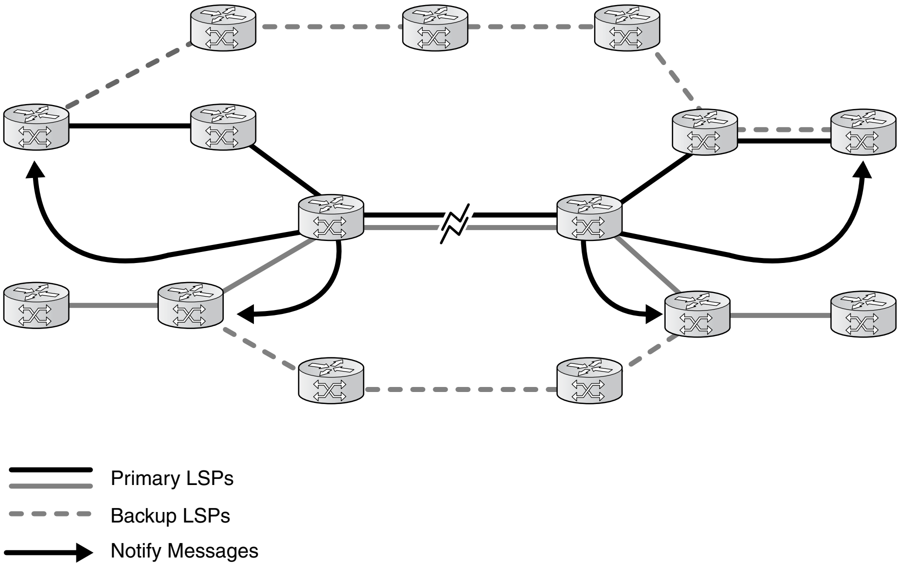


Figure 8-4
Protection switching with Notify messages.

- A PathErr message is intercepted and (minimally) processed at each LSR that it passes through, which impacts the time taken to notify the initiator of the problem.
- There may be many affected LSPs, all of which need PathErr messages to be sent at the same time—this can clog the network and consume important resources at the LSRs.

It is possible that an upstream node receiving the PathErr could incorrectly decide to take responsibility for the error, try to initiate local repair itself, and not forward the PathErr upstream until this had failed (or not sent any message at all). This could introduce extremely large delays into the propagation of the error to a repair point.

GMPLS introduces a new message to RSVP-TE only to enhance this notification process. The Notify message is sent directly from the point of failure detection to the point of repair. In Figure 8-4, when the black LSP fails, a Notify message is sent direct to the initiator. This message is sent directly to the initiator and does not have the IP Router Alert flag set, so it is not intercepted by transit nodes that are MPLS capable. It may optionally use IP double encapsulation. It uses regular IP routing to take the best route to the initiator.

If the LSP is bidirectional, a Notify message is also sent to the terminator to enable it to switch its data flow to the backup path. This message is sent by the node that detects the error on the terminator side of the failed link.

In fact, there is no reason why the repair points should be limited to the initiator and terminator LSRs. In Figure 8-4, the gray LSP is protected by a shorter backup segment shown by a dashed line. An LSR that wants to be told about LSP failure adds an object containing its own address to the Path or Resv message as it passes through. This object specifically requests the use of Notify messages and identifies the recipient of the message. The Path message identifies the Notify recipient on the initiator side, and the Resv indicates the Notify recipient on the terminator side.

As the Path or Resv flows through the network, the LSRs may change the Notify recipient and may even change whether or not a Notify is requested. This enables multiple repair points to be defined for an LSP.

Finally, it is possible that the most efficient use of Notify is to send it to a network device that can recalculate a path for the LSP and communicate it to the initiator using a management protocol (such as the Simple Network Management Protocol [SNMP]). In this case, the Notify recipient might be entirely distinct from the LSRs that make up the LSP.

Reporting Multiple Failed LSPs As described previously, one of the principle concerns with using PathErr to report LSP failures is that a single link failure may impact many LSPs, causing the need to send one PathErr message for each LSP. This could cause a data spike on the control channel and might cause temporary resource shortages in the LSRs that process the messages.

For this reason, the Notify message is defined so that it can report multiple LSPs at once. The only constraints are as follows:

- All failed LSPs on a single Notify message must have failed with the same reason code. Only one error spec is present and refers to all the reported LSPs.
- All failed LSPs on a single Notify must have the same Notify recipient (repair point).

State Removal on PathErr RSVP specifies that errors are notified to upstream nodes using PathErr messages and to downstream nodes using ResvErr messages. When the ingress receives a PathErr, it may

- Ignore the error, enabling state refresh to repair any damage.
- Reissue a Path message with different parameters that have a better chance of success.
- Tear the reservation by sending PathTear.

This approach is somewhat inefficient after a serious failure because two message flows are needed: the PathErr flow and the PathTear flow.

GMPLS introduces an option on a PathErr message to report that state has been removed at the sending node. That is, the flag indicates that the LSR that sent the PathErr has deprogrammed the switch for the LSP and has released all control blocks for the LSP.

The State Removed flag is applied on a hop-by-hop basis. When an LSR receives a PathErr with the flag set, it must send a PathErr upstream. It may choose whether to set the flag on the PathErr that it sends. If it does not set the flag on the PathErr it sends, it may also choose to store the fact that it received the flag from the downstream LSR.

Figure 8-5 shows the flow of messages for a failed LSP. In the example, the node that detects the error sends a Notify message to the ingress and sends a PathErr upstream indicating that the state has been removed. The PathErr propagates upstream with the next two nodes and does not set the State Removed flag. Farther upstream, an LSR decides to reenable the

Figure 8-5
Tearing state with PathErr.

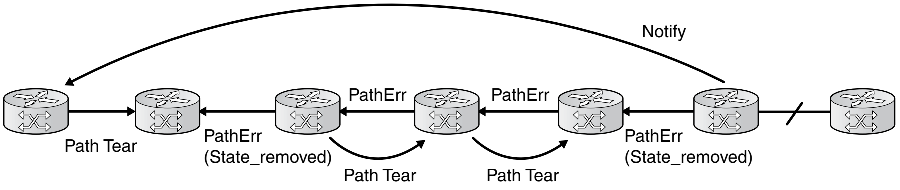
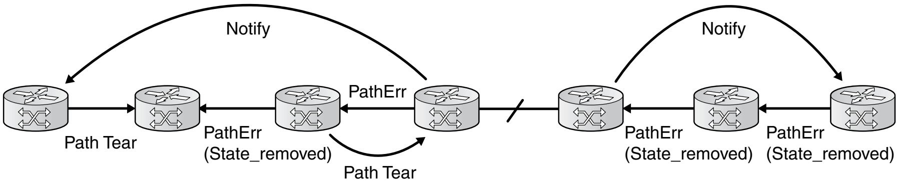


Figure 8-6
Tidying up in a fragmented network.



State Removed flag and so also sends a PathTear downstream. Just one node away from the ingress, the PathErr traveling upstream meets a PathTear traveling downstream and no further processing is required.

This feature can be useful to tidy up when the network is fragmented (see Figure 8-6). Upstream of the LSP failure on the left-hand side of the diagram, the processing is much like that described for Figure 8-5. The combination of PathErr (with or without the State Removed flag set) and PathTear (triggered by the Notify or the PathErr) causes the removal of state in a timely manner.

Downstream of the break, however, RSVP only provides for state removal through Path refresh timeouts. The LSR that detects the error might send Notify or ResvErr, and the egress might send ResvTear, but these exchanges do not fully remove the state or reservations at the intermediate nodes. Further, the PathTear sent by the ingress will never reach the LSRs downstream of the LSP failure if that failure affects the control path as well as the data path.

This situation can be remedied by using the State Removed flag on the PathErr message (see Figure 8-6). When it receives the Notify message, the egress LSR can remove state, deprogram its switch, and release reservations. It then sends the PathErr upstream, indicating that it has removed state, and the fragmented LSP is tidied up.

8.5 Bidirectionality

Trunks through the core of an optical network often need to be bidirectional. That is, they need to be capable of carrying data in both directions.

In the original specification of MPLS, bidirectional connections required the setup of two unidirectional LSPs and therefore coordination between the endpoints. This may have required management messages to be sent to both ends of the LSP, requesting that the two directions be signaled by the two ingress nodes. This had the disadvantage of needing two signaling protocols (one for MPLS and one for the management messages) and could result in the two directions of the LSP following different paths through the network. It also left the endpoints with the need to invent a way of identifying and coordinating the two unidirectional LSPs that made up a single bidirectional LSP.

An improvement was achieved using special processing at the egress of a unidirectional LSP to react to the receipt of a request for an LSP not just by sending a response, but also by sending a request in the opposite direction. With the use of route recording and explicit routing, this allowed the two directions to follow the same path through the network, but still left coordination problems between the two directions of LSP.

Both of these solutions have the additional issue that four signaling messages are required (request and response in each direction) to set up the single LSP.

GMPLS makes extensions to MPLS to address all of these issues and to allow a bidirectional LSP to be set up using a single message exchange. This has the benefit of requiring less signaling and achieves immediate coordination between the directions of flow.

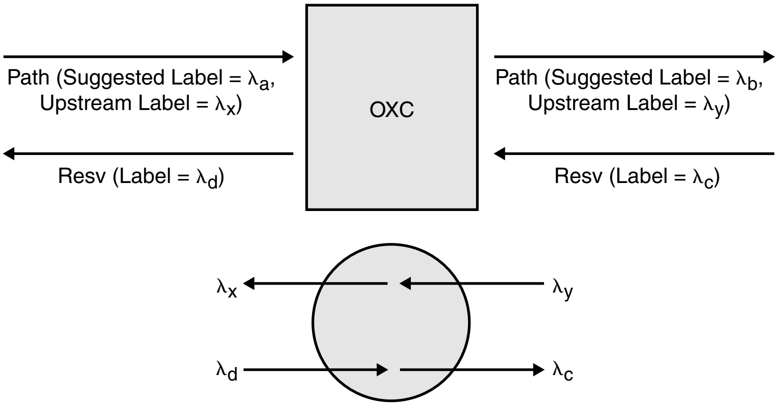
For the sake of clarity, we define the head end of the LSP as the *initiator* and the tail end as the *terminator*. In a unidirectional LSP, the ingress is the initiator and the egress is the terminator. In a bidirectional LSP, the concepts of ingress and egress and upstream and downstream are not clear because data flows in both directions, but as the LSP is originally requested from one place and responded to from another, we can use the terms “initiator” and “terminator” without ambiguity.

8.5.1 Upstream Labels

An upstream label is a new object introduced to LSP Setup Requests by GMPLS. It enables an upstream LSR to signal the label that should be used by the adjacent downstream LSR to forward data in the direction from terminator to initiator. Figure 8-7 shows the exchange of signaling messages at an individual LSR. Again, RSVP is used as the example protocol:

- A Path message is received at the LSR. As described in a previous section, this message carries a suggested label, which is used to preprogram the switch for the forward data path.
- The Path message also carries an upstream label. If the indicated label (λx) is not suitable for use at the LSR, it must reject the Path message (PathErr) at once. In this case, the PathErr message may include an acceptable Label Set object to tell the upstream LSR which labels would have been successful. The upstream LSR may retry the Path message using a different upstream label chosen from the acceptable label set or may fail the LSP setup by propagating the PathErr farther upstream.
- If the upstream label value is acceptable, the LSR selects a label (λy) to use as an upstream label on the outgoing interface and signals this in the upstream label of the Path message it sends out. If the switch is not capable of label modification, it will map the incoming upstream label value to the outgoing one (λy=λx).
- When the Path message reaches the LSP terminator and the switch there has been programmed, the reverse data path is complete and may be used immediately.
- The forward data path still has to be confirmed or reprogrammed as described previously, and the Resv message is used to carry the forward

Figure 8-7 Bidirectional LSP setup.



path labels (λc and λd) back toward the initiator. When the Resv passes through the LSR, the switch ends up programmed, as illustrated in the figure.

- Note that there is no mention of the upstream label on the Resv.

8.5.2 Confirming the Forward Path

One concern with the process described previously is that the GMPLS specifications state that when the LSP Setup Request (Path message) reaches the terminator, the reverse data path must be considered to have been established and data can immediately start to flow.

Delay at the Terminator Node In order to guarantee this, each LSR on the path might consider that it has to wait until its switch is fully programmed before it forwards the LSP Setup Request toward the terminator. This would result in slow LSP setup as each switch may take a considerable amount of time to stabilize. This defeats the purpose of label suggestion.

Alternatively, the LSRs on the path could simply forward the LSP Setup Request as soon as they have checked that programming the switch is likely to succeed and continue with the programming in the background. This relies on several procedures:

- Prior to sending data, the terminator LSR must leave a small amount of time for other LSRs to become successfully programmed. This time might be no longer than it takes to program its own switch, but it could be longer in a network made up of switches from different vendors.
- Programming errors at transit LSRs must be immediately reported as LSP failures, and the terminator LSR must be prepared to handle these error notifications even if they arrive ahead of the original programming request.

This approach enables the LSP setup time to converge on the unidirectional LSP setup time described previously, but it requires shared knowledge and cooperation between the switch implementations at the transit LSRs.

ResvConf A more secure solution is available in RSVP-TE and uses the ResvConf message. This message flows in the same direction as a Path message and confirms the receipt at the initiator of the Resv message. The terminator requests it by a flag on the Resv message.

The terminator can use this message to confirm that the switches at the transit nodes have been successfully programmed. In fact, the process would be that the Resv is not propagated toward the initiator by any LSR until it is happy that its switch is correctly programmed. When the Resv reaches the initiator, it sends a ResvConf to the terminator (hop by hop) and data can flow in both directions.

This solution increases the number of signaling exchanges (three rather than two), but increases the stability of the signaling. The choice of this option is entirely up to the implementation of the terminator LSR and requires no additions to the RSVP protocol. This option is not available in CR-LDP.

8.5.3 The Need for Asymmetry

Given that GMPLS supports both unidirectional initiator-to-terminator and bidirectional LSPs, it is natural to consider whether unidirectional reverse-path LSPs, that is, for data flow from the terminator back toward the initiator, would also be useful.

In fact, they are useful. In a densely used network, populated mostly with bidirectional (symmetric) and unidirectional forward LSPs, it may not be possible to find a route for a new bidirectional data flow that has enough available resources for both LSP directions. In this scenario, the only solution may be to signal the forward and reverse halves of the desired LSP separately. The resulting combination is known as an *asymmetric* bidirectional LSP.

Asymmetric bidirectional LSPs can be set up using any of the pre-GMPLS methods that were used to set up symmetric bidirectional LSPs. In fact, one of the drawbacks of these methods, the difficulty in getting the two directions to follow the same path, is not an issue for asymmetric bidirectional LSPs. However, the issue of needing a management and signaling protocol still applies, and the problem of coordinating the forward and reverse halves of the LSP still exists.

Reverse-path LSP signaling might also be achieved by reusing the GMPLS support for bidirectional LSPs with the (not yet standardized) understanding that as long as an upstream label is present, the Label Request and Label objects may be omitted from Path and Resv messages, respectively. Given this understanding, an asymmetric bidirectional LSP, with the two LSP directions following different paths, can be set up by separately signaling the forward and reverse halves of the LSP. Note that practical network administration tools would tie the two halves of the asymmetric LSP together such that they can be managed as a single logical entity.

8.6 Link Management Protocol (LMP)

When GMPLS is used to signal LSPs across core optical networks, a number of new management issues arise:

- If most of the switches in an optical network are photonic, and therefore may not automatically spot loss of light (LoL), how can the precise location of a fault be localized?
- If the link between two switches consists of a large bundle of optical fibers, how can the routing protocol be protected from having to advertise this huge number of links?
- With a large number of fibers, how can the neighboring devices agree on how to address these links without manually configuring each node with the other's link-numbering scheme?
- Given that some channel session protocol is required to address these issues, what other features of this protocol are needed, for example, to authenticate session peers and detect interrupted sessions?

LMP[3] has evolved to address these issues. This section presents some of its features.

Between two nodes in an optical network, there may be multiple parallel optical data links that carry the data traffic. These data links may be

- Electrically terminated at a node, in which case the node is always aware of the flow of data
- Transparent, for example, if a node is implemented using Micro-Electro Mechanical Systems (MEMS) mirror technology, in which case the node cannot normally see the data flowing (although in many optical devices, an optical tap can be used to check the data flow when necessary)

In the general case, as described previously in Section 8.3, there is a separate control channel that is used for protocol exchange (out of band), although it is also possible for the protocol exchange to take place through one or more of the data links (in band). The fundamental job of LMP is to validate the "wiring" of the links between adjacent nodes, validate that each data link is operational, and localize faults. Therefore, LMP protocol exchange is only required between adjacent nodes that are directly connected by data links.

The IETF LMP specification covers the following areas of functionality. Some of these areas are optional within the protocol and do not need to be present in an LMP implementation. Each area is described in greater detail in the following subsections:

- **Control channel management** This covers the establishment, configuration, and maintenance of an IP control channel between a pair of neighboring LMP nodes.
- **Link verification** This covers the verification of the connectivity of data links, together with dynamic determination of the mapping between local and remote interface IDs. If link verification is not supported, these mappings must be provided by configuration.
- **Link property correlation** This is confirmation between neighboring nodes that the mappings between local and remote interface IDs and the aggregation of multiple data links into traffic-engineering links are consistent.
- **Fault management** Lightpaths typically traverse multiple data links going from ingress to egress. When this lightpath fails, LMP provides a way to localize which data link has failed.
- **Authentication** This provides cryptographic confirmation of the identity of the neighboring node.

8.6.1 Control Channel Management

Control channel management is concerned with the negotiation and maintenance of the LMP session itself. LMP Hello messages are used to confirm that the session is still running. Parameter negotiation is necessary to agree on values for per-session parameters such as Hello timeout periods, supported protocol levels, and support for optional protocol features:

- **Control channel selection** This is not currently within the scope of LMP, but nevertheless is an important requirement for LMP to work. LMP peers must know to use the same control channel(s) when they exchange LMP messages.

- **Parameter negotiation** This enables a pair of LMP nodes to converge on a mutually acceptable set of configuration parameters. The protocol exchange involves three messages (Config, ConfigAck, and ConfigNack), each of which can contain a variable number of subobjects describing particular parameters (such as timeout values and support for optional protocol elements).

  The results of the parameter negotiation process are an agreement on timeout intervals for control channel maintenance and an agreement on the supported set of LMP capabilities.
- **Control channel maintenance** This involves an ongoing exchange of Hello messages between a pair of LMP nodes to confirm that the control channel is still operational.

  Control channel failure can result in the transfer of LMP operations to a backup control channel. If no LMP control channel is up, it is impossible to be sure of the state of the data links and LMP cannot collect any further information. The data links are marked as "degraded" and the signaling components can be warned that problems may occur and not be properly rectified. The signaling components could take remedial action such as rerouting the LSPs.

8.6.2 Traffic-Engineering Links

There may be a large number of optical fibers between a pair of optical switches. Since it is usually the case that several of these fibers have the same properties as far as routing protocols are concerned—that is, the same endpoints, QoS, and so on—it makes sense not to flood the routing database with duplicated information for each individual fiber. For routing purposes, therefore, fibers with the same properties are grouped together into a traffic-engineering link, and only the traffic-engineering link is advertised via routing.

However, if individual fibers are not known through routing information, GMPLS LSRs need a way of identifying individual fibers within a traffic-engineering link so that a signal switched into a fiber by the upstream LSR can be picked up by the downstream one. LMP meets this need by providing mechanisms for

- Mapping the ID used at one end of an optical fiber to the ID used at its other end (link verification)
- Grouping sets of IDs at either end into the corresponding traffic-engineering link ID (link summarization)

8.6.3 Link Verification

Link verification has two purposes:

- The verification of data connectivity on particular data links
- The automatic determination of the mapping between local and remote interface IDs for both data and traffic-engineering links

In GMPLS, LSRs use the interface ID mappings determined by LMP link verification to signal exactly which fiber of a traffic-engineering link is the target for an LSP.

The link verification process involves a number of steps:

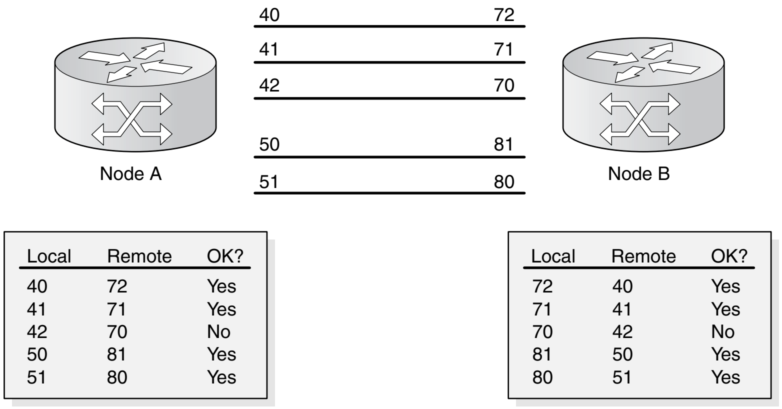
1. The exchange of BeginVerify/BeginVerifyAck/BeginVerifyNack messages between the LMP nodes to initiate link verification. The BeginVerify message that is sent by the initiating node includes the ID of the traffic-engineering link whose fibers it intends to test.
2. The sending of Test messages down the links to be tested one after another, together with timer/retry processing. Note that this message is transmitted on the particular data link being verified, *not* sent via the normal IP control channel. The receiving node checks for light on each fiber that it believes makes up its end of the specified traffic-engineering link.
3. The exchange of TestStatusFail/TestStatusSuccess/TestStatusAck messages to report the results of data link tests.
4. The exchange of EndVerify/EndVerifyAck.

The results of the link verification process, as illustrated in Figure 8-8, are

- The determination of which data links have tested successfully
- The mappings between the local and remote interface IDs for the data links
- The mappings between the local and remote interface IDs for the traffic-engineering links

8.6.4 Link Property Summarization

The main purpose of this function is to discover and agree between two adjacent LMP nodes on interface ID mappings and data link properties.



Local	Remote	OK?
40	72	Yes
41	71	Yes
42	70	No
50	81	Yes
51	80	Yes

Local	Remote	OK?
72	40	Yes
71	41	Yes
70	42	No
81	50	Yes
80	51	Yes

Figure 8-8
Link verification table.

This is achieved by the exchange of LinkSummary/LinkSummaryAck/ LinkSummaryNack messages, each of which contains multiple subobjects. These subobjects include

- Traffic-engineering link subobjects indicate the mappings between local and remote traffic-engineering link IDs, together with other information about the traffic-engineering links (such as protection and multiplex capability). Each traffic-engineering link aggregates multiple data links, which is indicated by the subsequent data link subobjects.
- Data link subobjects indicate the mappings between local and remote interface IDs (as determined by link verification or by manual configuration) along with information about the link type.

The results of the link property correlation process are

- Confirmation that the mappings between local and remote interface IDs are consistent between the LMP nodes (which is particularly important where the mappings have been preconfigured rather than dynamically determined via link verification)
- Confirmation that the aggregation of data links into traffic-engineering links is consistent between the adjacent LMP nodes
- Agreement on the properties and capabilities of the data channels

The aggregation of data links into traffic-engineering links is illustrated in Figure 8-9.

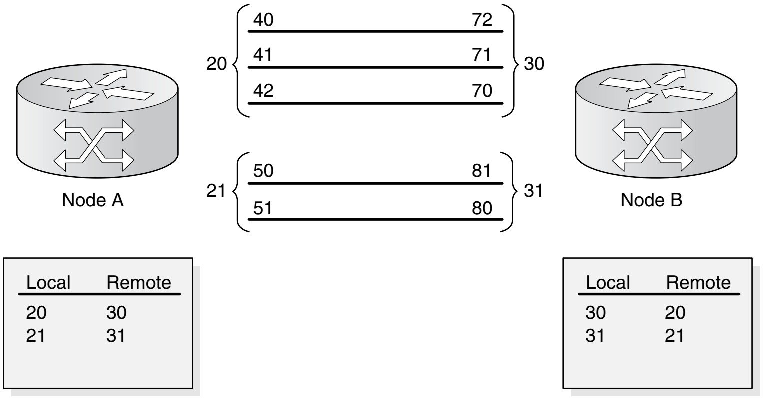


Figure 8-9 Traffic-engineering link mappings.

8.6.5 Fault Detection

Many photonic optical switches are transparent in the sense that they propagate the light signal without any interference. They can switch data by fiber, wavelength, or time slot without needing to examine the actual signal at all. Consequently, if the signal disappears because of a fault somewhere upstream, the switch may not even notice.

In the typical worst case, the absence of the signal is only detected when it needs to be converted back to electronic form for forwarding over a packet-switched network. Then, the only information is that there is a fault on one of the optical links somewhere upstream from the LSP through which packets were expected to arrive.

LMP fault detection is used in this scenario to localize the link on which the fault occurred. It proceeds as follows:

1. The process is initiated by the downstream node, which first notices failure on a data link.
2. This node sends a ChannelFailure message to its upstream neighbor for that link.
3. The upstream node determines which incoming/upstream data link(s) connects to the failed outgoing/downstream data link.
4. If the corresponding incoming data link has also failed, then the upstream node responds with a ChannelFailureAck and propagates the fault detection process by sending a new ChannelFailure message to the next node upstream. Alternatively, this node might have noticed

the LoL under its own stream, so it might have already propagated the ChannelFailure upstream.

5. If the corresponding incoming data link has not failed, then the upstream node has localized the fault, responds to the downstream node with a ChannelFailureNack message, and reports the error locally.

The results of the process are

- An indication that a particular data link has failed
- An indication of whether the failure has been localized to this node—precisely, to the downstream link at this node

LMP also includes a mechanism for nodes to indicate when particular data links recover again via the exchange of ChannelActive/ChannelActiveAck messages.

8.6.6 Authentication

Because the control channel between a pair of LMP nodes may pass through an arbitrary IP cloud (even the Internet), it is important to be able to authenticate the messages that are received over this channel.

To see why this is necessary, consider the denial of service (DoS) that could be caused by sending a fake LMP BeginVerify message to a core optical switch to instruct it to begin verification of all of its data links. If the link verification proceeded, normal data forwarding would be suspended for the duration of the testing.

Optionally, therefore, LMP sessions may negotiate the use of an MD5 hash algorithm to authenticate control messages. When authentication is agreed, each control message is signed by adding to it a 16-byte signature calculated from a shared secret key and an MD5 hash of the message contents.

8.7 Optical User-Network Interface (O-UNI)

The GMPLS extensions provide a way to use MPLS for provisioning in optical networks. In a real-world network, optical-networking equipment is typically used at the center of the network, providing a high-capacity, fast-backbone service for client devices that use existing low-cost and low-bandwidth networking technologies (for example, IP routers or ATM switches). This section seeks to answer the following question: how do the packet or MPLS networks request services from the time slot or lambda-switched optical transport network? There are two main possibilities, with others acting as hybrids between the two: the peer model and the overlay model.

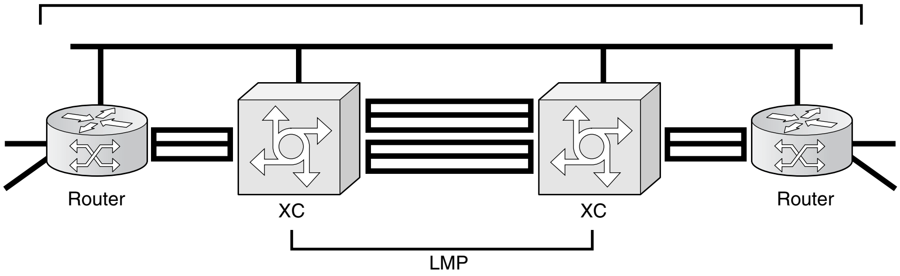
8.7.1 Models

Peer Model In the peer model shown in Figure 8-10, GMPLS is used to set up the whole path, including both the packet and transport elements, and the client's requests are spliced into the requests needed to set up the internal tunnel. Although this has the advantage of providing a single management layer to the service provider, there are some significant considerations to make:

- The service provider may want to keep topology information of the optical core private.
- Much of the equipment available for the optical core today uses proprietary signaling rather than GMPLS.
- Service providers may want to keep the paths set in the transport core as stable as possible and managed on a overall network basis for protection and restoration purposes. Thus, they would not want the service requests of the packet network to be fulfilled automatically.

Overlay Model In the overlay model (see Figure 8-11), the optical and packet networks are managed independently and may potentially be signaled by different protocols. One solution is for the Optical Transport

Figure 8-10
GMPLS peer model.



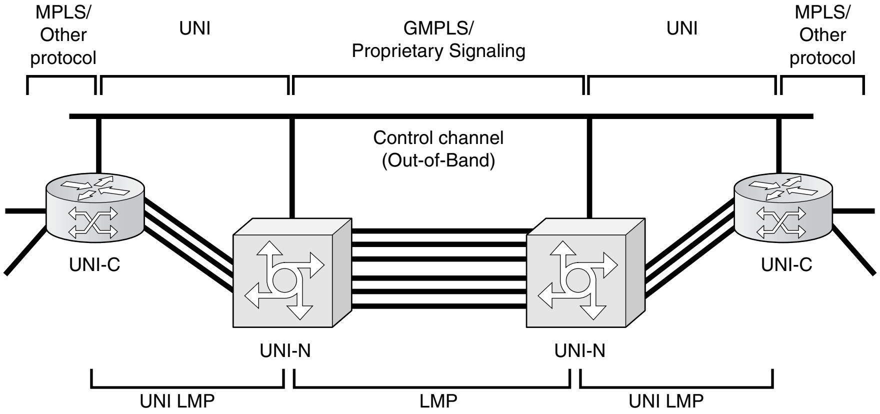


Figure 8-11 Optical Internetworking Forum (OIF) UNI overlay model.

Network (OTN) to provide a UNI that enables client network devices to request connections across it dynamically.

A signaling protocol needs to be defined to bridge the UNI (see Figure 8-12).

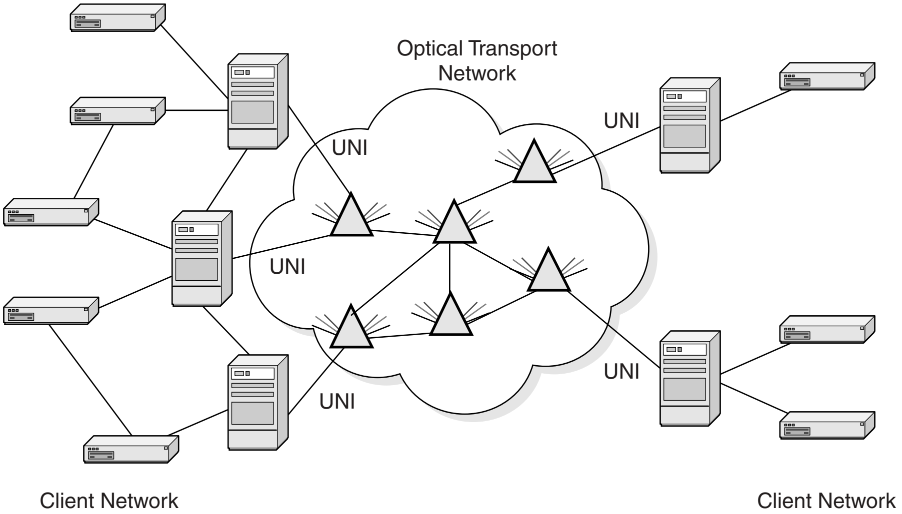
An internal tunnel between endpoint UNI-N nodes is set up using either GMPLS or some proprietary signaling mechanism. This tunnel can be provisioned ahead of time or on demand in response to the client request. The data path between the client nodes is then set up by overlaying it upon this internal tunnel. For the remainder of this section, we will look at the requirements and specifications of this O-UNI.

8.7.2 UNI Services

The client and OTN edge nodes that support UNI are not traditional signaling peers in that the OTN service provider agrees on a contract (in the billable sense) with the client to provide a certain level of service across the network. A UNI message from the client to the OTN is best viewed as a request for service.

Given the contractual angle, it is important to have a defined method for identifying and authenticating clients across the UNI. This enables the OTN to ensure that a client's account is in good standing before granting a connection request. For billing and auditing purposes, the OTN may also need to keep track of the connection by assigning a connection identifier that is valid beyond the lifetime of the connection and by returning this identifier to the client.

Figure 8-12
The OTN and UNI access.



In order to set up a connection, a client needs to discover what services are available from the OTN and then signal its requirements across the UNI. The types of service parameters signaled across the UNI are

- Requested bandwidth of the connection
- Class of service (CoS) (for example, the network restoration/protection strategy required)
- Diversity (discussed in the following section)
- Data-plane-specific characteristics (for example, for SONET/SDH, port, transparency, and concatenation information)

A fundamental requirement of UNI is that clients are not given access to OTN internal addresses or topology information. This means that UNI connection requests are not allowed to specify explicit routes. The diversity parameter enables a UNI client to request that a new connection must follow a different route to a set of previous connections without requiring internal knowledge of the OTN. This is necessary so that the client can request a disjoint backup to a primary route.

8.7.3 Addressing and Routing

A client requesting a connection across an OTN needs to identify the endpoints of the connection—the source and destination—using some addressing scheme.

The O-UNI solution is supposed to use a new, distinct address space for UNI clients. In this strategy, the OTN assigns a single, globally unique address to each client endpoint. We refer to this type of address as a transport network assigned (TNA) address.

Using an addressing scheme that is decoupled from both internal OTN and client protocol addresses has the following benefits:

- Internal OTN addresses are not revealed across administrative boundaries. Instead, clients address each other using TNA addresses, and mapping between the assigned address and internal switch addresses is done within the bounds of the OTN. (The alternative—revealing internal OTN addresses outside the OTN boundary—is viewed as a significant security risk.)
- The OTN can organize the TNA address to maximize the efficiency of routing and reachability calculations for its clients.
- By decoupling TNAs from client addresses, UNI can provide a seamless way for clients using different protocols to interwork. At the same time, the use of TNAs imposes certain responsibilities on the clients and OTN.
    - Clients must register with the OTN to be assigned a TNA.
    - The OTN must provide an address resolution service, which translates between a remote client protocol address and a remote TNA address.
    - The OTN must be capable of distributing TNA address client mappings and reachability information to all edge nodes that provide the UNI (intradomain).
    - Different OTNs must be capable of distributing addressing information between them (interdomain).

One rejected addressing strategy was to use client protocol addresses. The key disadvantage of this strategy is that clients of the OTN are likely to use a variety of different address spaces and formats (such as Network Service Access Point [NSAP], IPv4, IPv6, and e.164), which leads to a number of difficulties:

- The OTN has no control over these address spaces so it cannot organize them to facilitate routing (for example, via hierarchies and summarization).
- Clients who need to interconnect to other clients using different protocols would need to be capable of translating between multiple address formats.

- Within a single address format, two client networks may use addresses that clash (for example, in the case of virtual private networks [VPNs]). This means that a client address on its own may not actually be a unique identifier.

8.7.4 The Realization of UNI in GMPLS Protocols

Although UNI is not a standard GMPLS peer-to-peer relationship, the signaling function required by UNI is very close to that achieved by the existing GMPLS signaling protocols RSVP and CR-LDP (although in practice, as with GMPLS, RSVP is well ahead in terms of current implementation). With a few extensions to support the UNI-specific features such as TNA addressing, diversity, and connection identification, these protocols can be used for UNI signaling.

Similarly, the various link management and service discovery functions carried out by client and OTN UNI neighbors can be achieved by an extended flavor of LMP (see Section 8.6).

The mapping between the standard and UNI flavors of the protocols are nontrivial so an edge node for a GMPLS-provisioned OTN needs to be capable of running UNI across some interfaces and standard optical signaling across others.

The OIF has produced a specification document describing how existing protocols can be extended to provide UNI functionality. The OIF UNI currently covers SONET/SDH connections and includes a wide range of protocol features discussed in this chapter, including the following:

- RSVP, LDP, and LMP with UNI extensions
- UNI TNA addressing
- Out-of-band signaling
- Bidirectional LSPs
- Label set
- Reduced refresh processing
- Setup confirmation

8.7.5 LMP Extensions

In addition to the standard LMP features described in Section 8.6, OIF UNI defines two LMP extensions:

- **In-band control channel bootstrap** This is a way of initiating in-band control channel communication when the IP address of a neighboring node is not known. The idea is that when a control channel is realized in band over a data link, an OIF UNI LMP node can send its initial LMP Config message to the All Nodes multicast address rather than to a specific unicast IP address. This enables an LMP node to automatically determine which IP address to use when communicating with its neighbor.
- **Service discovery** This is a way for UNI client and network nodes to discover each other's UNI capabilities via the exchange of ServiceConfig, ServiceConfigAck, and ServiceConfigNack messages. Using this mechanism,
  - A UNI client node indicates its support for port-level attributes such as link type, signal type, and transparency.
  - A UNI network node indicates its support for particular signaling protocols, transparency and monitoring, and network diversity.

In summary, optical networking is an important and challenging domain for development and experimentation using the GMPLS protocols. Conversely, the use of GMPLS and LMP is a major step forward for optical network providers who have historically been troubled by interoperability problems resulting from nonstandardized control and management protocols.

The application of GMPLS to the optical network has thrown up a number of interesting technical challenges. This paper has described some of those challenges and presented explanations, examples of how those challenges can be addressed in practice, and, in particular, how aspects of the GMPLS protocol have evolved to meet those challenges.

End Notes

1. Peter Ashwood-Smith et al., "Generalized MPLS—Signaling Functional Description," *draft-ietf-mpls-generalized-signaling-07.txt*, November 2001.

2. This section was contributed by Data Connection Ltd. (DCL) based on Neil Jerram and Adrian Farrel's paper "MPLS in Optical Networks—An Analysis of the Features of MPLS and Generalized MPLS and Their Application to Optical Networks, with Reference to the Link Management Protocol and Optical UNI," Data Connection Ltd. White Paper, Version 2, www.dataconnection.com, October 2001. DCL is the leading independent developer and supplier of MPLS, ATM, SS7, MGCP/MEGACO, SSCTP, VoIP conferencing, messaging, directory, and SNA portable products. DCL's customers include Alcatel, Cabletron, Cisco, Fujitsu, Hewlett-Packard, Hitachi, IBM Corp., Microsoft, Nortel, SGI, and Sun. DCL is headquartered in London and has U.S. offices in Reston, Virginia, and Alameda, California. It was founded in 1981 and is privately held. In 1999, the company received its second Queen's Award for outstanding export performance.
3. "Link Management Protocol (LMP)," *draft-ietf-ccamp-lmp,* November 2001.
4. "Fault Tolerance for LDP and CR-LDP," *draft-ietf-mpls-ldp-ft*, October 2000.
5. RFC 2961, "RSVP Refresh Overhead Reduction Extensions," April 2001.
6. Data Connection Ltd., "Surviving Failures in MPLS Networks—An Examination of the Methods for Protecting MPLS LSPs Against Failures of Network Resources," White Paper, www.dataconnection.com, October 2001.

CHAPTER 9

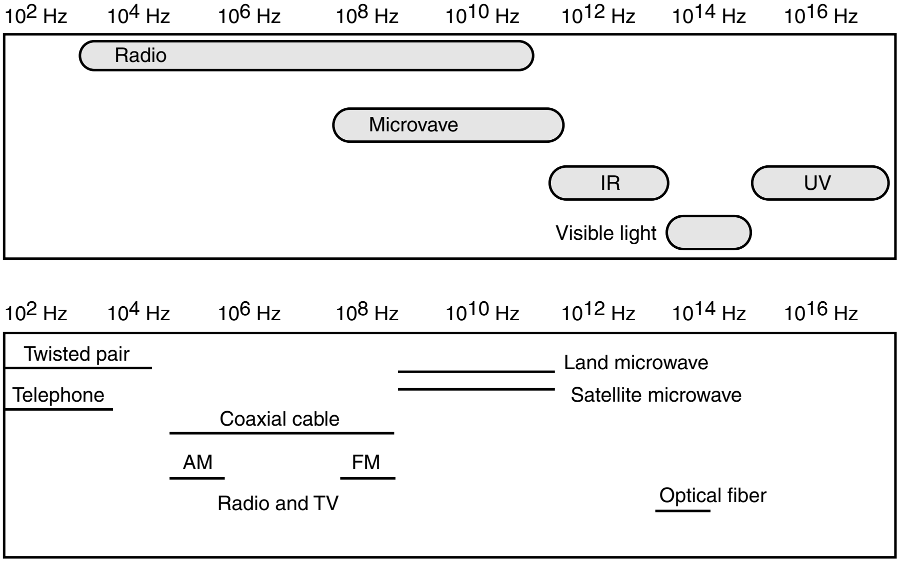
Free-Space Optics (FSO)

Using free-space optics (FSO) to bridge the last mile and first mile between a local area network (LAN), metropolitan area network (MAN), and wide area network (WAN) in a wireless environment has been seriously propounded as the solution for MANs. Considering the cost of connecting buildings to core optical networks running in the street below or a block or two away (typically in the range of $75,000 to $200,000 and at times more), FSO is an attractive contender. However, like all radio solutions proposed for this purpose in the past 25 years, starting with Xerox's original mid-1970s proposal, it appears to have limited or transient application. This chapter addresses the advantages and disadvantages of FSO and the challenges that FSO faces based on technical, financial, and implementation aspects. The chapter wraps up by analyzing the future of FSO.

The global telecommunications networking environment has undergone unprecedented growth over the past several years, which was fueled by the telecommunications deregulation in 1996.[1] Since 1996, service providers have spent an enormous amount of capital investment on building long-haul WAN infrastructures such as Frame Relay, Asynchronous Transfer Mode (ATM), and Internet Protocol (IP) networks. In the past two years, emphasis has been put on MANs due to the increasing telecommunications traffic demand in densely populated metropolitan cities. In order to enable end users (that is, corporate users) to access these long-haul WANs and MANs, network designers must provide them with a broad array of flexible access solutions that operate in a cost-effective manner. Even today, a vast majority of local-loop connections (sometimes referred to as *first mile* or *last mile*) are still limited to T1 line speed, and consequently, there is a definite need for utilizing an access technology that will bridge the gap between a LAN and MAN as well as between a LAN and WAN. Optical wireless systems (FSOs) represent an approach for addressing and possibly solving the first-mile and last-mile bottleneck. Figure 9-1 depicts the bandwidth capabilities of various transmission media.

Optical wireless systems transport high-speed data through the air on light beams transmitted by highly focused lasers. These lasers are often installed on building rooftops and windows. This emerging technology has now matured to a stage where a noticeable group of equipment vendors have introduced their own FSO products for various applications. Some of the most promising features for using FSO as an access solution include the ease of network deployment, low startup and operating costs, fiber-like bandwidth, and no requirement for an operating license. Commercially available systems can offer capacities ranging from 100 Mbps to OC-48 (2.5 Gbps) and beta systems operating at data rates as high as 160 Gbps.

Figure 9-1 *Bandwidth capability of various media.*



Note: In general one can get between 2 and 10 bits per baud (hertz) of digital throughput.

Note: Maximum bandwidth range for FSO: Terabits per second (10^{12} bps, or 1 Tbps); actual throughput is significantly lower.

9.1 Technical Aspects of FSO

9.1.1 Network Architecture

The network architecture of an optical wireless (FSO) system is categorized by one of the four network topologies that follow:

- Point-to-point topology
- Point-to-multipoint topology
- Meshed topology
- Ring topology

In a point-to-point FSO system, a transmitting laser that sits on a rooftop aims at a receiving laser within its line of sight, and an optical data transport link is set up between the transmitter and the receiver through the air. In ideal atmospheric conditions, this configuration enables data to reach between 2 to 4 km with a data rate between 155 and 622 Mbps. Two of the advantages of using this configuration are that there is dedicated bandwidth between two endpoints and no complex network architecture or planning requirements. However, two of the disadvantages are that the optical link is the single point of failure and service availability lowers as the distance between the two endpoints increases. Figure 9-2 diagrams a point-to-point topology.

Often referred to as a *hub-and-spoke architecture*, a point-to-multipoint FSO system consists of a transceiving laser that acts as a hub for serving and aggregating optical signals sent from multiple lasers being served by the hub. Because the bandwidth is shared among multiple users (lasers), each laser is able to burst up to higher rates if bandwidth is available. This type of configuration can support data rates between 155 and 622 Mbps with a predefined distance of 1 to 2 km among the hub and its serving lasers. A key advantage of this topology is the efficient connection to the core network because all traffic is collected at a single point (hub). The point-to-multipoint topology shares some of the same disadvantages as the point-to-point topology. An additional disadvantage is that the development and setup costs for establishing the hub are high. Figure 9-3 depicts point-to-multipoint configuration.

Figure 9-2
Point-to-point free-space (infrared) link.

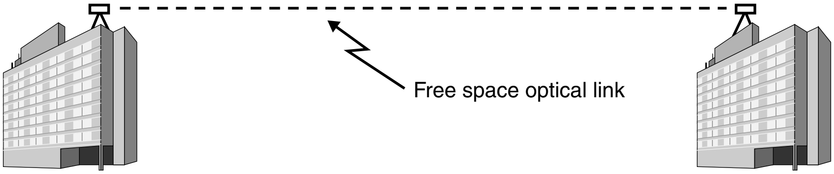
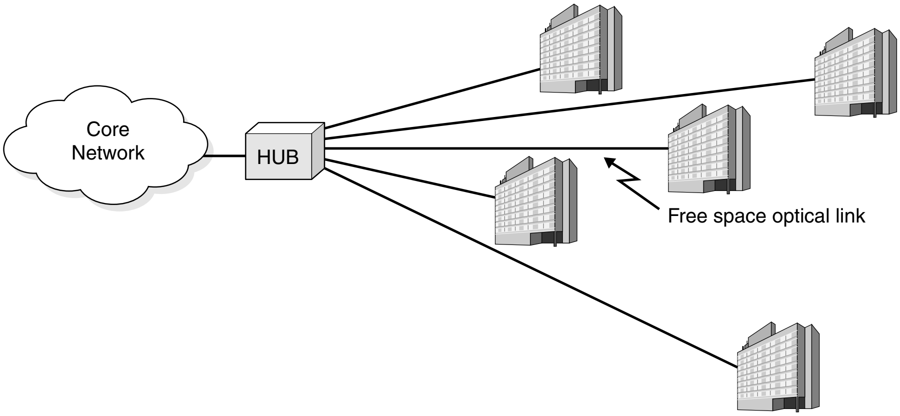


Figure 9-3
Point-to-multipoint configuration.



The third network topology in an optical wireless system is a meshed architecture. This topology is sometimes considered an extension of the point-to-multipoint topology. In a meshed FSO system, all transceiving lasers are interconnected or linked together. In other words, rather than linking an individual laser to a hub in the hub-and-spoke topology, this same laser is connected to all other lasers installed in the system in a meshed topology. This type of configuration can also support data rates between 155 and 622 Mbps with a distance limitation of 200 to 450 meters between two lasers. Two key advantages of utilizing this topology are that it offers higher availability from shorter optical links and diverse paths between two endpoints. However, some disadvantages exist, such as the requirement of a large number of over-the-air hops or multiple overlapping rings and the complex requirements for network design and planning. Figure 9-4 shows a meshed FSO system configuration.[2]

The fourth network topology for an FSO system is called a ring architecture. Analogous to fiber-optic cable networks that exist today, a ring topology uses the same concept except it is done wirelessly. An arbitrary transceiving laser is only configured to communicate with its two neighboring lasers. This laser is not set up to exchange data with all other lasers within the network as it is in a meshed architecture. A key advantage in this topology is that there is no single point of failure. The ring architecture has the same disadvantages as the meshed architecture. Ring topology is shown in Figure 9-5.

After you've seen the four types of network topology in an optical wireless system, it is useful to see an example of how carriers apply them in a densely

Figure 9-4
Meshed FSO system configuration.

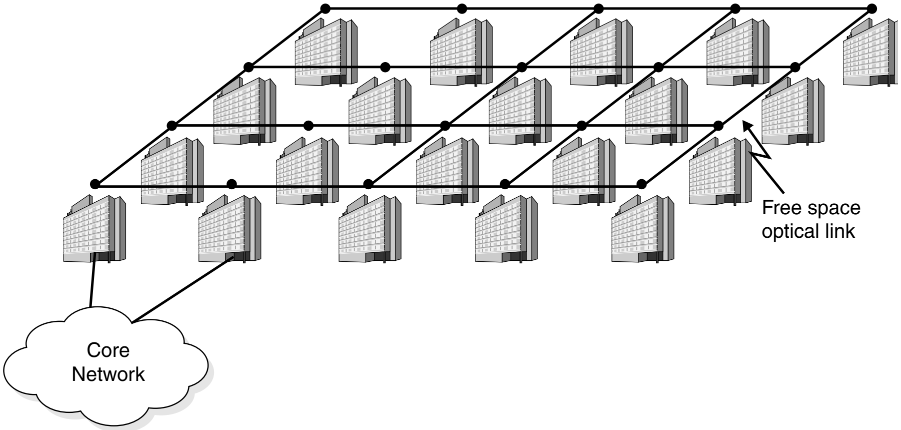
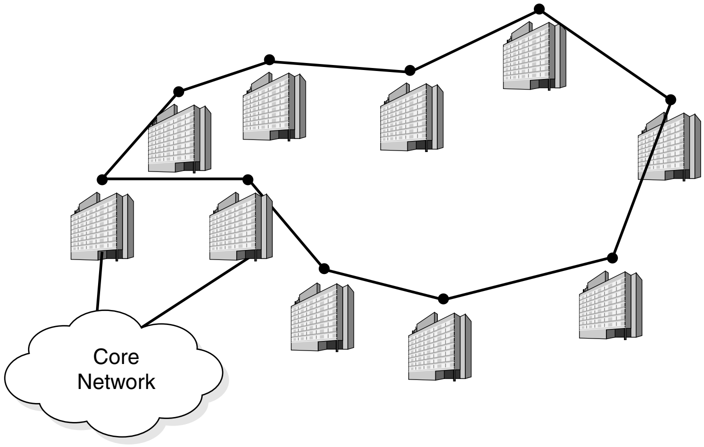


Figure 9-5
Ring FSO topology.

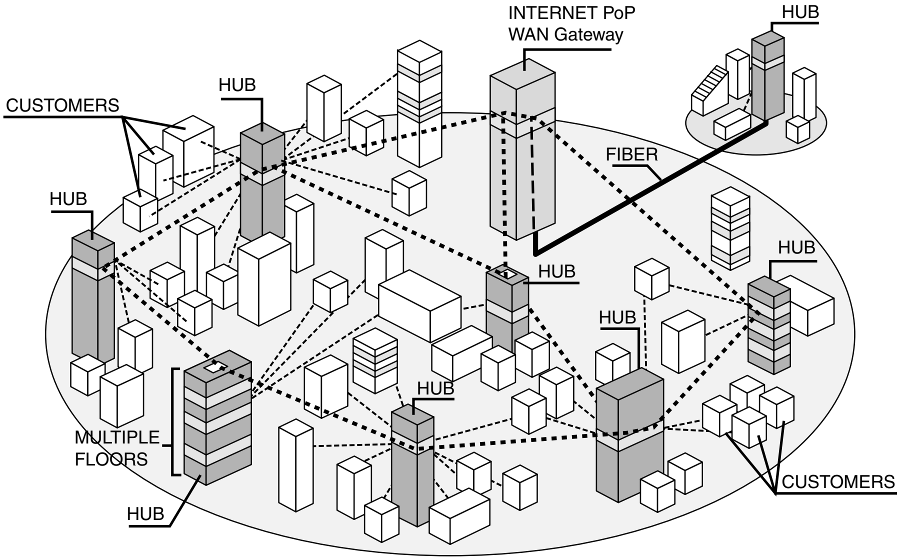


populated high-rise office building environment, or more specifically, in a MAN environment. In Figure 9-6,[3] the tall light grey building represents a carrier's point of presence (POP) that serves as a gateway between a WAN and MAN. The MAN is established via an optical wireless (FSO) system. The carrier's POP exchanges information with its nearby hubs (dark grey buildings) via back-haul FSO links. In this example, these back-haul FSO links are configured in a combined ring and meshed topology. Communication between the carrier's hub and its serving customer premise buildings (white buildings connected by thin lines to the hub buildings) is set up in a hub-and-spoke or point-to-multipoint topology. Figure 9-6 is an example of how these network topologies can be applied to the MAN environment.

9.1.2 Key Network Elements

An FSO system consists of two major network elements: free-space optical telescopic transceivers and network management software and hardware. The free-space optical transceiver acts as a transmitting and receiving terminal for transporting optical signals to another transceiver. For the transceivers to interface properly with an area network such as a LAN and MAN, network management software and hardware are necessary. The hardware would be a switch that is usually bundled with the transceiver from the same manufacturer. The network management software controls how the optical signals are transmitted and received from the transceiver and how the signals are routed to an area network. If high availability is a crucial

Figure 9-6 Example of MAN environment.



factor when utilizing the FSO system, a high-speed microwave radio backup link also needs to be installed. In order to understand how FSO network elements work in detail, it is beneficial to describe the design aspect of an optical transceiver.

The design of an optical transceiver can be categorized into two types: reflective optical design and directive optical design. In the case of the reflective optical design, the goal is to couple incident light onto the core of a multimode fiber that is connected to a switch. Each transceiver consists of an individual receiver aperture as well as four individual transmitting apertures. The receiver aperture is usually connected to a multimode fiber with a FC-PC jumper. The four individual transmitting apertures are linked to single-mode fiber jumpers, while an equal length of fiber is used to provide the optical signal to each transmitting aperture. In addition, two 785-nm lasers are used as alignment beacons and are integrated into the transceiver housing for initial alignment purposes. Figure 9-7 illustrates the design of an optical transceiver.[4]

The second optical transceiver design utilizes a convergent lens to focus laser beams onto a light-sensing element; thus, it is called a directive optical design. The transmitter portion of this design is composed of the fiber-optic interface, interface circuitry, laser diode, and beam collimator. The fiber-optic interface is an optical transceiver module operating with data rates from 100 to 622 Mbps using multimode fiber with distances up to

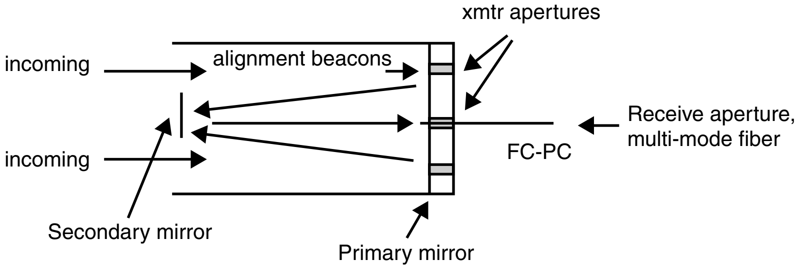


Figure 9-7
Design of an optical transceiver.

1,100 meters. The optical interface output is coupled to the laser diode driver circuitry. The laser diode driver circuitry provides the threshold bias and modulation current for the laser diode. The laser diode is a commercially available near-infrared laser diode operating at a wavelength of 850 nm. The driver circuit is configured to provide an output power of 10 mW from the laser diode and focused through a collimator lens for transmission.

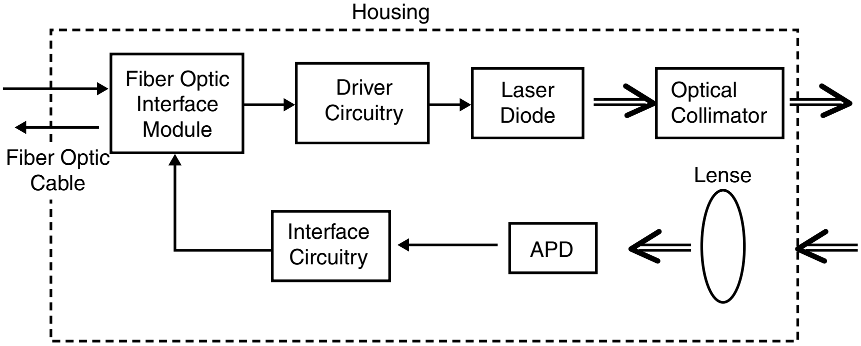
The receiver portion of the transceiver uses a convergent lens for the collection and focusing of the beam onto the optical-sensing element such as the avalanche photodetector diode (APD). The output from the optical-sensing element (typically APD) is transmitted through the interface circuitry to the optical transceiver module, which is then transmitted out on the optical fiber to the network. The block diagram in Figure 9-8 depicts how the modules interwork.[5]

Having described the two designs that are typically used to construct an FSO optical transceiver, it is a good idea to discuss briefly the hardware and software required to interoperate with the optical transceiver. The hardware is usually a router and/or switch that is capable of accepting fiber handoffs from the output of an optical transceiver. For controlling the interoperability between a transceiver and a router and/or switch, a network management interface is necessary. In most cases, this interface is based on the Simple Network Management Protocol (SNMP) and provides additional features such as network monitoring, a graphical user interface (GUI), and remote management capabilities.

9.1.3 Atmospheric Impairments and Other Problematic Issues

Weather Attenuation The most significant difference between optical wireless (FSO) and fiber-optic laser transmission is the unpredictability

Figure 9-8
Block diagram of typical system.



of the attenuation of laser power in the atmosphere compared to fiber.[6] Optical fiber cables attenuate at a consistent, predictable rate. Typical measurements of this attenuation range from 2 to 3 dB/km for multimode fiber cables and 0.2 to 0.5 dB/km in the case of single-mode fiber cables. In contrast, the atmospheric attenuation of laser power in optical wireless transmission can vary dramatically and is difficult to predict. Atmospheric attenuation loss spans from 0.2 dB/km in exceptionally clear weather to 350 dB/km in a very dense fog. The main reason behind this phenomenon is that the moisture (fog) particles are so tiny and dense that they act as millions of tiny prisms, and when a band of light shines through them, the signal is distorted and dissipated. The effect of fog on FSO atmospheric attenuation is comparatively more severe than rain. This is because the fog droplet radius (5 to 20 μm) is on the order of the laser wavelengths that results in noticeable attenuation (up to 350 dB/km) compared to the attenuation (20 dB/km) introduced by the rain droplet with sizes between 200 and 2,000 μm. Therefore, the effect of weather attenuation on FSO laser power is considered one of the major obstacles for deployment because it reduces the uptime or availability of FSO communication systems.

In order to explain the underlying reason for atmospheric attenuation in further detail, it is essential to state the two types of atmospheric impairment in FSO systems that cannot be neglected. The first type of impairment is described as *scintillation*, which by definition is the intensity fluctuation in the received optical signal. Scintillation causes error bursts during periods of atmospheric turbulence that is originated by temperature gradients near the ground,[4] resulting in localized random pockets of varying indices of refraction. When the size of these scintillation pockets is smaller than the diameter of the laser beam in an optical transceiver, the optical beam experiences distortion, which is observable as a nonuniform optical intensity across the wave front (see Figure 9-9).[4]

On the other hand, when the size of the pockets is larger than the diameter of the laser beam, the beam will randomly get off course of the site path. This effect is described as *beam wander* (see Figure 9-10).[4]

The mixing of both scintillation and beam wander leads to the fluctuations in overall signal stability. Therefore, the optical beam experiences an in-homogeneous spatial and temporal propagation path, causing the beam to propagate in a random fashion. When combined together, both effects produce an overall noise component to the received optical signal. Atmospheric attenuation is the dominant cause of error bursts in free-space communications systems and it becomes very important to find ways to characterize and quantify it.

Moving Buildings The second physical issue that affects the performance and reliability of an FSO system is the natural movement of buildings. Although it is rarely noticeable by humans, high-rise buildings occasionally experience minimal oscillation, or *jitter*, due to the movement or vibration of objects such as a heavy-duty electric generator. This creates a potential problem when using optical transceivers because the transmitters and receivers must be aligned accurately to each other in order to minimize transmission error.

Flying Objects Flying object obstruction is the third potential issue that can deteriorate the performance of an FSO system. The most significant instance is the flying pigeon or bird scenario. For network topologies such as the point-to-point or hub-and-spoke configuration, this issue can introduce a potential problem because these topologies only provide a single

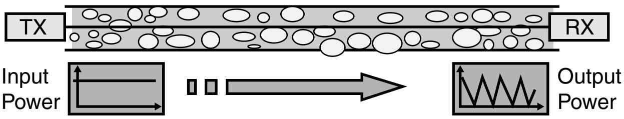


Figure 9-9
Scintillation.

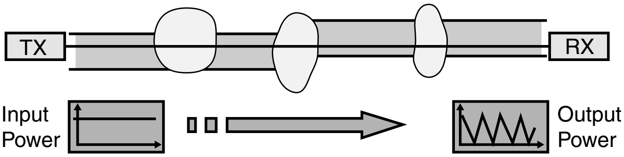


Figure 9-10
Beam wander.

path for transmission. However, for a meshed or a ring topology, this issue would not affect the availability of the FSO system because there is an alternate path to route the signal in these two topologies.

Eye Safety Almost all wireless optical laser transceivers from various FSO vendors currently operate at a wavelength between 785 and 1,550 nm. This particular mode of operation used by most laser transceivers draws a major concern of the laser ocular hazard introduced to the human eye. This is because the tissues of the human eye are very susceptible to damage that is dependent on wavelength, power, and pulse duration. More specifically, specific tissues in the eye interact strongly with different portions of the optical spectrum. The cornea and lens are transparent to the visible wavelengths and near-infrared exposure as they are intended to focus visible light onto the retina. However, mid-infrared (1,400 nm to 3 μm) to far-infrared (3 μm to 1 mm) exposure to the cornea and lens can cause photokeratitus and cataracts.[7] Due to the potential danger of a transceiver laser to the human eye, it is critical to enforce stringent laser safety standards established earlier to all optical transceiver products distributed worldwide. One of the widely adopted laser safety standards comes from the International Electrotechnical Commission (IEC). The IEC 60825-1 *Safety of Laser Products Part 1—Equipment Classification*, *Requirements*, *and User Guide* is an internationally recognized technical standard for users and manufacturers of laser products. This standard applies to the laser safety classification system that is mandatory on all wireless optical laser transceivers. Table 9-1[7] shows the classification system of the IEC 60825-1, as applied to the example of a doubled CW Nd:YAG laser.

Table 9-1

Classification System of the IEC 60825-1

Class	Description	AEL
1	Lasers that are safe under reasonably foreseeable conditions.	0.39 μW
2	Visible lasers (400–700 nm) where aversion response such as the blink reflex affords eye protection.	1 mW
3a	Lasers that are safe for viewing with the unaided eye. Direct viewing of the beam with aids (such as binoculars) may be hazardous.	5 mW
3b	Lasers that are hazardous when the beam is viewed directly.	500 mW
4	Lasers capable of producing hazardous diffuse reflections and that may pose skin and fire hazards.	>500 mW

A safe operation of the laser transceiver requires human access to be limited to Class 1 radiation. A wireless optical network deployed in a space where a casual pass by could intercept a laser beam requires the system to be Class 1 and incapable of delivering radiation in excess of Class 1. In order to determine the classification of an FSO system, design features such as the number and size of transmitted beams, divergence, power, wavelength, pulse duration, physical size of terminals, and installation location must be considered. In addition, active safety systems that control the laser in response to a potential access event are sometimes implemented. In summary, laser safety in a wireless optical network is a key parameter that all existing and future laser transceivers must comply with.

9.1.4 Major Vendors

From the time when FSO technology was in the development phase until it became commercially available, it has been recognized that approximately 16 wireless optics companies have introduced their own version of FSO solutions that are targeted for different market segments. The goal of this chapter is not to explain the products of these 16 companies in great detail, but rather to point out some of the key players in this emerging technology.

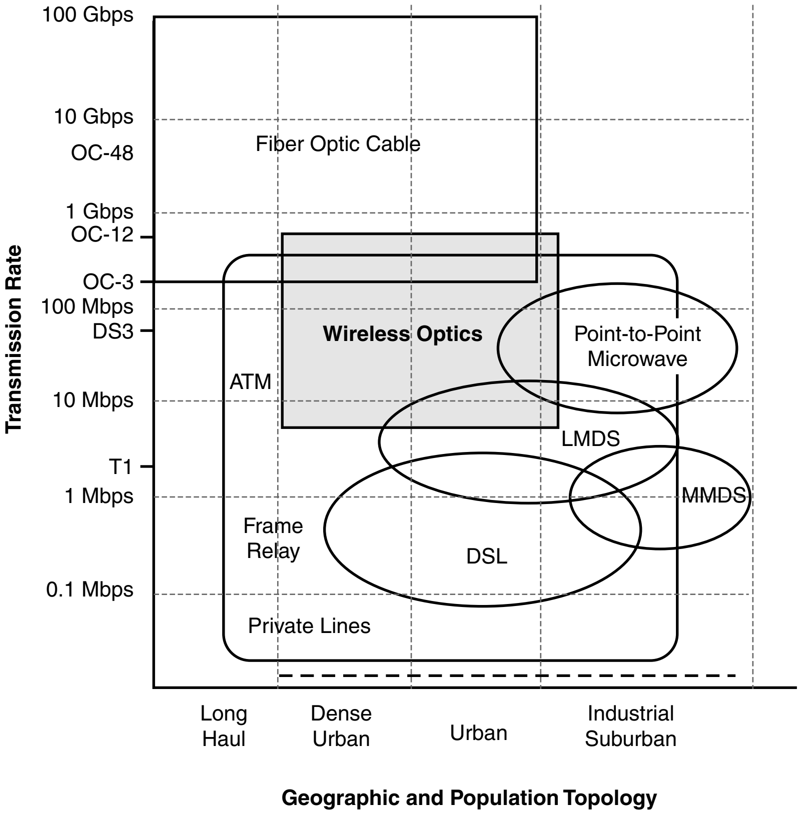
AirFiber's ATM-based OptiMesh product targets facility-based carriers looking to extend the reach of their existing fiber networks. Its product can be used in point-to-point, point-to-multipoint, or redundant mesh configurations and features shorter links for improved availability. OptiMesh contains two major network elements. The first element is the Roof-Top System (RTS), or node, which is equipped with two or four optical transceivers, an ATM switch, a drop to the building demarcation, a control microprocessor, and software. The second element is the Element Management System (EMS), a graphical, user-friendly, crossplatform software package that manages the mesh network and each of its nodes. LightPointe Communications' Flight product overcomes the availability issue with "five 9s" (99.999 percent) reliability by using a backup hybrid optical/microwave link. Its solution will fit within any carrier network protocol or topology, including mesh, point-to-point, ring, and point-to-multipoint configurations. Flight consists of three critical elements. The first element is FlightTransport, which is an optical transceiving equipment unit. The second element is FlightManager, which is a robust SNMP-based network management interface between the FSO link heads and the customer's network operations center (NOC). The third element is FlightNavigator, which is a tool set that permits deployment simplicity and has the capability to perform network traffic-monitoring functions such as traffic simulation and link availability analysis. Optical Access Inc. has introduced a carrier-class Optical Access Mesh, which features a built-in radio frequency (RF) wireless backup system to handle potential downtime caused by transmission attenuation. It supports Fast Ethernet and Gigabit Ethernet (GbE) and includes switching and routing technology that enables bandwidth distribution and provisioning for IP networks. Optical Access Mesh contains Terescope (laser transceiver), OptiSwitch, SNMP Access, and MegaVision (network management system). A Canadian-based FSO manufacturer called fSONA introduced the SONAbeam product, which transmits at speeds ranging from 155 Mbps to 1.25 Gbps and can be configured in different topologies such as point-to-point and inter-LAN. This product also has a radio-based wireless backup to tackle against weather interference.

9.2 Financial Aspect of FSO

9.2.1 Comparison Between FSO and Other Established Access Technologies

Among the first-mile and last-mile access technologies that are available and well established today, some of these technologies are more suitable for certain geographical environments and/or populations. The most well-known access technology that is utilized by both residential and commercial users in an urban or rural area is offered via copper cables (twisted pairs). However, in certain remote locations, you might consider using wireless local loop or satellite communications as an alternative access loop. These access technologies do not serve the purpose of offering broadband access to commercial customers located in densely populated metropolitan areas. This is where Digital Subscriber Line (DSL) comes into play. DSL is supposed to provide a connection that is higher than T1 speed in a cost-effective way, but the recent downfall of DSL service providers and equipment manufacturers has dramatically hindered the growth of this access technology. Some customers have been considering wireless access technologies such as microwave, Local Multipoint Distribution Service (LMDS), and Multichannel Multipoint Distribution System (MMDS). These are well-designed access technologies despite the fact that they are not widely deployed by service providers and the performance could vary significantly from location to location. Fiber-optic cables can offer a tremendous amount of bandwidth, but the question is whether customers need the amount of bandwidth and are willing to pay the cost of fiber installation and the systems required. In addition, the buildout of a brand-new fiber system can take six months to a year even before the customer can benefit from the fiber network. FSOs fit nicely into the arena from a cost and lead-time standpoint. FSO is targeted for customers who require quick provisioning time and high bandwidth that ranges from Ethernet speed (10 Mbps) to OC-3 or OC-12. FSO is an ideal candidate if customers are located in sites near a service provider's backbone. Figure 9-11 shows how the technologies mentioned previously address different market environments (that is, urban, rural, and remote) based on the transmission rates supported.[2]

Figure 9-11 *Addressing different market environments.*

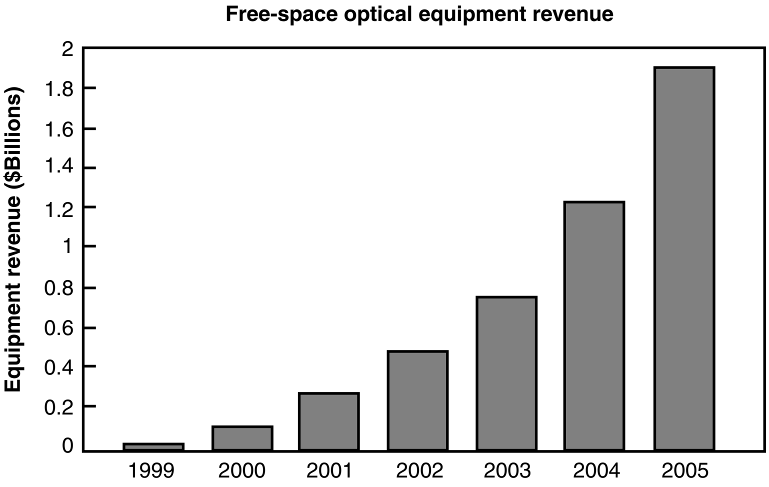


9.2.2 Global FSO Equipment Revenue Forecast

By the Strategis Group According to a telecommunications technology market study report by the Strategis Group, FSO is expected to become a strong player in the telecommunications industry over the next five years. They projected that worldwide FSO equipment revenues are expected to increase from under \$100 million in 2000 to \$2 billion by 2005.[8] Figure 9-12 charts the revenue growth projection for FSO equipment between 1999 to 2005.

By Allied Business Intelligence In addition to the FSO market study report done by the Strategis Group, a market researcher based in New York called Allied Business Intelligence completed a similar report to demonstrate the revenue projection of FSO until 2005. They projected that the global FSO transceiver revenue will reach more than \$4 billion by 2005, which doubles the revenue forecast in the FSO study report from the Strategis Group.[9]

Figure 9-12
Revenue growth projection for FSO equipment. Source: Strategis Group



9.3 Deployment of FSO

Given the fact that no spectrum license for deploying an FSO system is required, typically an optical wireless system with a point-to-point topology can be installed and brought into operation in approximately one hour or less. However, this is only true when a suitable line of sight is identified quickly.

9.3.1 Implementation of FSO Systems

In order to ensure a smooth and quick implementation of an FSO system, it is advantageous to follow the five-phase process depicted in the following list:

First phase	Site Acquisition and Zoning
Second phase	Site Survey
Third phase	Site Preparation
Fourth phase	Commissioning
Fifth phase	System Test and Acceptance

The Site Acquisition and Zoning (first) phase includes determining the proper rooftop locations for placing laser transceivers and obtaining building-owner approval for a rooftop installation. The Site Survey (second) phase is needed to ensure that each proposed rooftop location meets all network requirements (such as line of sight) and is evaluated for feasibility and accessibility. The purpose of the Site Preparation (third) phase is to prepare architectural and engineering drawings and specifications, obtain construction permits, install power to the rooftop location, and install tripod or mount support for the transceivers. The purpose of the Commissioning (fourth) phase, a critical phase in the entire deployment process, is to commission a new transceiver unit into the network. The Commissioning phase includes powering up the equipment, verifying the software, establishing wireless optical links with other transceivers within the network, setting up a network management channel back to the NOC, and the verifying proper operation of the wireless optical links. The System Test and Acceptance (fifth) phase contains the verification of system configuration and end-to-end functionality as well as system performance.

9.4 Strengths and Weaknesses of FSO

The following list summarizes some of the key strengths of FSO:

- Significantly lower cost for deploying an FSO network on average compared to the buildout of a new fiber optical solution.
- An FSO network can be implemented in a few days to a week; however, it typically takes several months to a year to install new optical fibers.
- Bandwidth can easily be scaled with a broad range of head room per link (from 10 Mbps to 1.25 Gbps).
- FSO bypasses license complexities compared to RF wireless technologies.
- FSO is network-friendly in the way that it can be engineered to be protocol independent and designed to be compatible with common monitoring and management protocols.

However, the weaknesses of FSO are

- The performance of FSO is very susceptible to weather attenuation, making this technology an alternative rather than primary solution for carrying mission-critical and sensitive data traffic.
- The general public's discomfort of using laser technology raises skepticism because of the potential hazard introduced by the use of optical wireless transceivers.
- Very narrow market segment (such as the demand in dense urban areas) makes it very difficult to compete with a fiber-based optical solution in metropolitan areas.

9.5 The Future of FSO

Whether a new telecommunications technology will be successful or not relies heavily on certain perspectives. FSO is clearly an emerging technology that requires the time and effort of both service providers and equipment manufacturers to promote as well as to educate their potential customers on this technology. When evaluating the future of FSO, it is a good idea to analyze it from two perspectives: demand demography and economy.

9.5.1 Demand Demography

The demand for enormous bandwidth has always originated from large-enterprise customers and in most cases these customers are situated in densely populated metropolitan cities. Most of these customers might have their own optical fiber network (dark fiber or dim fiber) and they might look for an alternative solution in the event that there is a fiber outage. FSO would be the perfect candidate for such applications because it can offer OC-x-equivalent bandwidth. An FSO network can be deployed in a week or two. However, FSO can rarely be considered the primary solution for these large customers, and this significantly hinders its demand growth.

9.5.2 Economy

The economy of a business environment can greatly affect the future of FSO. Back in early 1999, the e-commerce and Internet industries experienced explosive growth due to the tremedous bandwidth demand by both e-commerce and online companies. This growth triggered the rapid deployment of new fiber-optic systems across the United States. In contrast, toward the end of 2000 and early 2001, the failure of online companies and the general economic slowdown had left these new fiber-optic systems unused even through the latter part of 2001. Most enterprise customers even scaled back their traffic demand projection from an OC level to a DS3 or T1 level. This economic trend significantly impacted the demand of a broadband optical wireless solution such as FSO because these customers were left with a lot of fiber-based bandwidth and did not know what to do with it.

Similar to other emerging technologies, FSO can have a bright future due to its flexibility, scalability, and relative low cost. However, as pointed out earlier, the challenges that FSO faces could overcome all the promising attributes and eventually cause this technology to phase out.

End Notes

1. This chapter was first developed by Kenneth Wong in the summer of 2001 for Professor Minoli's graduate class at the Stevens Institute of Technology.
2. AirFiber Tutorial, "General Technology Overview," http://www.airfiber.com/products/faq/Html/helpset.htm.
3. Terabeam, "Terabeam: Our Technology: Last Mile," http://www.terabeam.com/our/our_net.shtml.
4. Paul F. Szajowski, Gerald Nykolak, James J. Auborn, Herman M. Presby, G. E. Tourgee, and Dennis M. Romain, "Key Elements of High-Speed WDM Terrestrial Free-Space Optical Communications Systems," *Free-Space Laser Communication Technologies XII, Proceedings of SPIE* 3932 (May 2000): 2–14.
5. Kenneth A. Morrison, "COTS Low-Cost 622 Mb/s Free-Space Laser Communications Link for Short-Distance Commercial Applications," *Free-Space Laser Communication Technologies XII*, *Proceedings of SPIE* 3932 (May 2000): 52–59.
6. Isaac I. Kim, "Statistical Study of Fog and Rain Applied to TereScope Fusion Availability," *Optical Access Incorporated*, (2000): http://www.freespaceoptic.com/WhitePaper_statistical_study_of_fog_and_rai.htm.
7. Jim Alwan, "Eye Safety and Wireless Optical Networks (WONs)," White Paper, AirFiber, Inc., (April 2001): http://www.airfiber.com/tech/whitepapers.htm.
8. "Free-Space Optics: Fixed Wireless Broadband," An industry study by the Strategis Group, October 2000.
9. Meghan Fuller, "Free-Space Optics No Longer Simply a Niche Solution," *Lightwave Magazine* (May 2001): 1, 50, 52.

References

Carbonneau, Theresa H. and G. Stephen Mecherle, "SONAbeam Optical Wireless Products." *Free-Space Laser Communication Technologies XII*, *Proceedings of SPIE* 3932 (May 2000): 45–51.

Heatley, David J. T., David R. Wisely, Ian Neild, and Peter Cochrane, "Optical Wireless—The Story So Far." *IEEE ComSoc Magazine* (December 1998): http://www.labs.bt.com/people/heatledj/papers/optical_wireless/optical_wireless.html.

Kim, Isaac I., Bruce McArthur, and Eric Korevaar, "Comparison of Laser Beam Propagation at 785 nm and 1,550 nm in Fog and Haze for Optical Wireless Communications." *Optical Access Incorporated*, (March 2001), http://www.freespaceoptic.com/WhitePaper_Comparison_of_Beam_in_Fog.htm.

Light Reading—The Global Site for Optical Networking, "Reports: Wireless Wonders—A Survey of Broadband Wireless Developments," (July 2000): http://www.lightreading.com/document.asp?doc_id=1207.

Masud, Sam, Sean Buckley, and Ted McKenna, "Ten Hottest Technologies—These Picks Sizzle, Sparkle, & Shine," *Telecommunications Magazine Online* (May 2001), http://www.telecommagazine.com/issues/200105/tcs/tenHottest.html.

Nykolak, Gerald, Paul F. Szajowski, A. Cashion, Herman M. Presby, G. E. Tourgee, and James J. Auborn, "40-Gb/s DWDM Free-Space Optical Transmission Link over 4.4 km." *Free-Space Laser Communication Technologies XII*, *Proceedings of SPIE* 3932 (May 2000): 16–20.

Rockwell, David and G. Stephen Mecherle, "Optical Wireless: Low-Cost Broadband Optical Access." *Lightwave Special Edition—Optical Carrier Solutions* (June 2001): 14–18.

CHAPTER 10

Practical Design and Economic Parameters

This chapter briefly explores the kind of parametric economic analysis needed to assess the financial situation of a greenfield carrier offering metro optical services. Established carriers need to take the issues discussed in this chapter into account and consider the entire embedded environment of networks and operations, inclusive of existing systems, rings, central offices (COs), customers, and so on. Recent history shows that the bottom line is most critical for carriers of all sorts.

We recently developed a spreadsheet model comprised of about 500,000 cells that supports the business case for 34 types of high-end broadband services (for example, 10-Mbps transparent LAN service [TLS], 100-Mbps TLS, and Gigabit Ethernet [GbE] TLS; access at 10, 100, and 1,000 Mbps; two-point private line at 10, 100, and 1,000 Mbps; access and private line at OC-x; Internet access with TLS or private line at OC-x; and so on) in 25 cities over a period of 5 years. This leads to a $1.2 billion cumulative revenue over the 5-year horizon ($650 million in year 5). An operation with solid 40 to 50 percent gross margins, 20 to 25 percent net margins, and earnings before interest, taxes, depreciation, and amortization (EBITDA) positive in two and a half years was demonstrated and launched. Quite a large number of factors need to be taken into account in a heavy-duty financial analysis. A subset of the input variables to such a model is shown in the following sections. We do not show the results—just the inputs—since the results are spread over three dozen very large worksheets.

10.1 Market Background

A startup carrier first needs to decide what kind of metro-level services it wants to offer. In order to be viable, a carrier needs to have a plethora of services (a couple of dozen or more). In addition to metro-level services, the carrier should also offer Internet services, long-haul services, and perhaps storage and disaster-recovery services. These, however, typically could be delivered in conjunction with partners to avoid the expense of building out a national Internet Protocol (IP) and general-transport network.

Carrier pricing needs to be sustainable and a profit must also be shown (at least within a couple of years). It is not desirable to sell service below cost as some Ethernet Local Exchange Carriers (ELECs) have done because this creates losses and a dependency on external funding. Eventually this leads to bankruptcy.

Metro-level providers need to concentrate on high-ticket-item services (in the range of $4,000 to $15,000 monthly recurring charge [MRC]) in order to be viable. If they concentrate at the low end (either sub-$1,000 or around $1,000 MRC), there will be less of an opportunity for margins, and the customer volume will be so high that the operations cost (to provision, engineer, and maintain service) will be rather large.

After a set of services is selected, your focus will turn to developing and deploying a multiarchitecture, multitechnology platform. Technology will have to be secured across multiple vendors to ensure best buys. Simultaneously, you'll need a market analysis that looks at the portfolio of services against potential customers in various cities (perhaps by focusing on specific Standard Industrial Classification [SIC] codes of interest, such as banking, insurance, brokerage, and so on). Results from your market analysis are necessary to understand the following:

- How many customers and/or orders can be secured in a city assuming some (conservative) low penetration rate (for example, 2 percent in year 1 growing to 5 percent in year 5)?
- How many customers and/or orders can be secured in a building? (Typically, this number is 3 to 5, not 20, 40, or 67 as some ELECs have assumed.)

The last two factors will determine the scope, kind, and architecture of the needed platform; they will also determine the price point to keep in mind.

Hence, look for the output of this stage of analysis to yield at least the following information:

- Kinds of services
- Orders per city per quarter per service type
- Target price points per service
- Total revenue per quarter

This information will drive the technical design.

10.2 Technical Design Parameters

Implicit in this step is also a rollout plan so that the right network is deployed at the right time to support the demand. (The "build it and they will come" model of network construction has been solidly refuted in past years.) In the following lists, we lay out a scenario to illustrate some of the variables essential for a sound business plan. The numbers shown are generalized and do not represent specific vendors or products.

The following list provides the key sets of inputs that comprise the modeling, such as the basic equipment costs, capital expenditures (capex), CO and in-building equipment costs, sales, general, and administrative (SGA), operations expenditures (opex), and so on.

The major parameters to consider include the following:

- Capex assumptions
- CO and building costs
- Facility and SGA employee expenses
- Operating assumptions
- Financing activities
- Revenue per employee
- City-level models (how many cities and when)
- Service revenues

The next list presents the kind of worksheets that will be derived in support of the analysis. The bottom of the list is the most discrete information; the top of the list is the most summarized information. The kind of output data sought by the modeling includes the following:

- Key financial information
- Balance sheet
- Cash flows
- Statement of operations
- Equity positions
- Revenues
- Interest income
- Building and CO order costs
- Operations and facility administration expenses
- Operations and facility administration per employee costs
- Operations and facility administration assumptions
- Other facilities
- SGA expenses
- SGA per employee costs
- Sales employees needed
- SGA assumptions
- Interest expense

- Fiber capitalization
- Equipment (fixed assets) summary
- Equipment and fiber cost calculation
- Order reaggregation to equipment
- Order fractionalization to cities

The key costs for an operation using a multiarchitecture, multitechnology design are identified in Table 10-1. These numbers are simply representative and do not correspond to any particular vendor or platform.

Table 10-1

Capex Assumptions

Item	Sub-element	Unit cost
Virtual LAN (VLAN)	Customer Premises Equipment (CPE) (with maintenance)	$17,738
	CO (with maintenance)	$160,646
Time-division multiplexing (TDM) (next-gen SONET)	CPE (with maintenance)	$38,210
	CO (with maintenance)	$91,708
Coarse wave division multiplexing/WDM (CWDM/WDM)	CPE (with maintenance)	$41,312
	CO (metro master) (with maintenance)	$44,920
Asynchronous Transfer Mode (ATM)	CPE (with maintenance)	$9,825
	CO (with maintenance)	$373,040
Private line	CPE	$28,712
	CO	$102,367
Storage and hosting	Servers and routers	$309,200
	Expansion modules	$52,750
Common equipment	CO (with maintenance)	$165,656
	PCs (apportioned per quarter)	$446
	Telemetry (per market)	$85,000
	Proportional network operations center (NOC) charge per market	$613,500
	Proportional lab charge per market	$500,000
	Management information (MIS) cost per market (LAN) system	$92,500
	Test equipment (ATM)	$85,000
	Test equipment (other)	$54,300
Operations support systems (OSSs)	OSSs cumulative through year 1	$1,800,000
	OSSs cumulative through year 2	$4,700,000
	OSSs cumulative through year 3	$9,000,000
	OSSs cumulative through year 4	$17,500,000
	OSSs cumulative through year 5	$25,000,000

Table 10-1 cont.

Capex Assumptions

Other assumptions			
Orders/building	3		
CPE per VLAN ring	16		
CPE per TDM ring	16		
Yearly interest for positive cash balances (one quarter hence)			4%
Yearly reduction (erosion) in capital costs			
	Equipment	Rings	
2nd year	95%	94%	
3rd year	90%	88%	
4th year	85%	82%	
5th year	80%	76%	
Financing arrangement: five years of capex financing with x% down			
	1st year	65%	
	2nd year	45%	
	3rd year	30%	
	4th year	20%	
	5th year	10%	
Capex financing rate		11%	

Opex costs are widely variable. Table 10-2 depicts some examples associated with services. For example, our operations number includes the allocated rent, power, and heating, ventilation, and air conditioning (HVAC) for the CO, which is typically a collocation facility. In addition, there may be crossconnect charges that the collocation facility may charge to the carrier. You also need to figure in your building costs, such as risers and payments to building owners.

Clearly, the carrier will have network costs. For metro applications, these include the metro access rings and the spurs or laterals into the buildings (see Table 10-3). Metro core rings will also cost, typically on a par with the access rings, as long as they are in the greater proximity of the core business district. Rings reaching into the suburbs will cost significantly more.

A major component of a carrier's expense is operations-related costs. Operations costs are derived in part from factors such as those shown in Table 10-4. Some item entries (such as sales salaries) will drive SGA.

Table 10-2

Assumptions for Costs for Orders at the CO and Building

	Common Components (per Order/Month)			
	Building Riser + Owner	**CO**	**Internet**	**WAN**
VLAN	200	300	300	400
TDM (next-gen SONET)	300	400	TBD	TBD
CWDM/WDM	300	400	TBD	TBD
Private line	200	300	TBD	TBD

Table 10-3

Facilities Costs

Item	**Sub-item**	**Quarterly costs (3×MRC)**	**Nonrecurring charge (NRC)**
VLAN			
	Rings	**$7,800**	**$62,000**
	Spurs	**$6,000**	**$5,000**
TDM (next-gen SONET)			
	Rings	$7,800	**$62,000**
	Spurs	**$6,000**	**$5,000**
CWDM/WDM			
	Rings	**$7,800**	**$62,000**
	Spurs	**$6,000**	**$5,000**
ATM			
	Point-to-point facilities	**$13,200**	**$3,600**

Note: Number of quarters for IRU (for example, 20, 40, or 80): 80

In order to achieve specific financial milestones, a carrier must scrupulously control SGA and operations cost. Even for a best-in-class carrier, SGA and operations costs tend to each represent 20 percent of the total budget. One approach is to look at the revenue you expect each employee to generate (in Table 10-5, an industry average of $290,000 per employee a year after a few years from inception). Revenue erosion also needs to be taken into account when calculating the number of emloyees:

	1st year	**2nd year**	**3rd year**	**4th year**	**5th year**
Revenue erosion, which drives added new revenue needs	0.02	0.02	0.02	0.02	0.02

Table 10-4

Facility and SGA Employee Expenses

Operations and Facility Administration	1st-year payroll	2nd-year payroll	3rd-year payroll	4th-year payroll	5th-year payroll
Engineering					
Manager	$90,000				
Engineer	TBD				
Associate	TBD				
Inflation rate	3.00%				
City Operations					
Manager	$100,000				
Supervisor	TBD				
Technician	TBD				
Inflation rate	3.00%				
Network Operation and Customer Care					
Manager	$75,000				
Supervisor	TBD				
Analyst	TBD				
Inflation rate	3.00%				
Network Planning					
Manager	$80,000				
Supervisor	TBD				
Associate	TBD				
Inflation rate	3.00%				
Technology					
Manager	TBD				
Engineer	TBD				
Associate	TBD				
Inflation rate	3.00%				
MIS					
Manager	TBD				
Programmers	TBD				
Associates	TBD				
Inflation rate	3.00%				

Table 10-4 cont.

Facility and SGA Employee Expenses

Operations and Facility Administration	1st-year payroll	2nd-year payroll	3rd-year payroll	4th-year payroll	5th-year payroll
SGA					
Administrative					
Executive	TBD				
Administrator	TBD				
Department manager	TBD				
Administrative assistant, bookkeeper, and clerk	TBD				
Inflation rate	3.00%				
Mktg./Prod. Mgmt./Pre and Post Sales Support					
Manager	TBD				
Supervisor	TBD				
Team member	TBD				
Inflation rate	3.00%				
Sales					
Regional manager	TBD	TBD	TBD	TBD	TBD
Branch manager	TBD	TBD	TBD	TBD	TBD
Account executive	TBD	TBD	TBD	TBD	TBD

Table 10-4 cont.

Facility and SGA Employee Expenses

SGA	1st-year payroll	2nd-year payroll	3rd-year payroll	4th-year payroll	5th-year payroll
		Legal			
Attorney	TBD	TBD	TBD	TBD	TBD
Inflation rate	3.00%				
Tax rate	9.00%				
Benefits	10.00%				
Insurance	0.27%				
Employee distribution					

Average Cost per Employee	Manager	Supervisor	Associate	Clerical	Total	
Engineering	0.0137	0.1096	0.8767	0.0000	1.0000	Span: 8 to 1 for all levels
In-city operations	0.0137	0.1096	0.8767	0.0000	1.0000	Span: 8 to 1 for all levels
NOC/Prov/Mtc	0.0137	0.1096	0.8767	0.0000	1.0000	Span: 8 to 1 for all levels
Network planning	0.0137	0.1096	0.8767	0.0000	1.0000	Span: 8 to 1 for all levels
Technology	0.0137	0.1096	0.8767	0.0000	1.0000	Span: 8 to 1 for all levels
MIS	0.0137	0.1096	0.8767	0.0000	1.0000	Span: 8 to 1 for all levels

Table 10-4 cont.

Facility and SGA Employee Expenses

Average Cost per Employee	Manager	Supervisor	Associate	Clerical	Total	
Administrative	0.0036	0.0073	0.0582	0.9309	1.0000	Span: 8 to 1 for all levels
Marketing and product management	0.0137	0.1096	0.8767	0.0000	1.0000	Span: 8 to 1 for all levels
Sales	0.0090	0.0910	0.9000	0.0000	1.0000	
Legal	1.0000	0.0000	0.0000	0.0000	1.0000	
			Facility and or SGA			

Support Costs	1st year	2nd year	3rd year	4th year	5th year
Customer meals and entertainment (based on employees)	TBD				
Financial services (fixed yearly amounts for banking, escrows, and so on)	TBD				
Money-raising fees, registration, and so on (as % total revenue)	20.000%	4.000%	1.500%	0.750%	0.375%
Insurance (based on quarterly revenue)	0.333%				
Legal, accounting, and professional					
Legal, accounting, and professional (fixed per quarter)	TBD	$15,000	$15,000	$15,000	$15,000
Legal, accounting, and professional (based on quarterly revenue)	0.250%				
Marketing software (fixed per quarter)	TBD				
Meals and entertainment (fixed)	TBD				

Table 10-4 cont.

Facility and SGA Employee Expenses

Support Costs	1st year	2nd year	3rd year	4th year	5th year
Meals and entertainment (incremental based on employees)	$180				
Miscellaneous (based on quarterly revenue)	0.21%	0.28%			
Miscellaneous (based on the number of cities)	TBD	TBD			
Miscellaneous (based on employees)	TBD	TBD			
Office expense (based on employees)	TBD	TBD			
Office supplies (based on employees)	TBD	TBD			
Paging services (based on employees)	TBD	TBD			
Promotion and advertising (fixed and incremental based on revenue)	TBD	TBD	4.000%	4.000%	4.000%
Rent (based on employees)	TBD	TBD			
Telephone					
Business (based on employees)	TBD	TBD			
Cellular (based on employees)	TBD	TBD			
Long distance (based on employees)	TBD	TBD			
Internet and WAN services (fixed per quarter)		TBD			
Training (based on employees)	TBD	TBD			
Travel and transportation expense (based on employees)	TBD	TBD			
Utilities (based on employees)	TBD	TBD			

Table 10-4 cont.

Facility and SGA Employee Expenses

Support Costs	1st year	2nd year	3rd year	4th year	5th year
Vehicle expense—facility and so on (based on employees)	$60				
Vehicle expense—sales (based on employees)	TBD				
Vehicle expense—other (based on employees)	TBD				
Vehicle lease expense (based on cities)	TBD				
Small tools (based on employees)	TBD				
Outsourced NOC expense (fixed amount)	$2,500,000	$250,000	$250,000	$0	$0

Table 10-5

Productivity Measures

	1st-year payroll	2nd-year payroll	3rd-year payroll	4th-year payroll	5th-year payroll
Revenue/employee	$70,000	$140,000	$280,000	$290,000	$290,000
Revenue/month/sales account executive	$8,000	$8,000	$9,000	$9,000	$10,000

Dividing the total carrier revenue by this goal yields the total number of employees in the company. Next, you need to establish a monthly quota for salespeople against the company revenue using the Rule of 78. The Rule of 78 yields the number of sales account representatives needed. Subtracting the salespeople from the total number of staffers just computed gives the number of employees that can be supported in each company division. (See Appendix A for a calculation of the normalized revenue per employee and Appendix B for a calculation of the number of salespeople required for revenue.)

Table 10-6 identifies the capital infusion, and timing thereof, that may be needed from financiers.

Table 10-6

Financing Measures

Financing Occurs		Shares	Monetary Amount	Share Value
		Financing Activities		
Year 1	1st qtr.	TBD	**$200,000,000**	TBD
	2nd qtr.	0	$0	TBD
	3rd qtr.	0	$0	TBD
	4th qtr.	0	$0	TBD
Year 2	1st qtr.	TBD	**$430,000,000**	TBD
	2nd qtr.	0	$0	TBD
	3rd qtr.	0	$0	TBD
	4th qtr.	0	$0	TBD

There are other relevant parameters to consider, such as the installation revenue, sales channels revenue percentages, and so on, which will affect SGA and salespeople counts. Some assumptions, but certainly not all, are captured in Table 10-7:

Table 10-7

Operating Assumptions

	1st year	2nd year	3rd year
Miscellaneous revenue from order installations (% collection)	15.00%	25.00%	35.00%
Amount per order	$3,500	$3,500	$3,500
Tax rate (single value)	39.00%		
Annual bad debt expense (based on revenues)	2.50%		

Distribution of revenue by sales elements	Fraction raw revenue	Fraction revenue retained	Fraction collected revenue	Fraction paidout revenue
1st year				
Retail, direct	0.60	1.00	1.00	0.00
Wholesale, direct	0.15	0.80	0.80	0.00
Retail, channel	0.15	0.90	1.00	0.10
Wholesale, channel	0.10	0.70	0.80	0.10

Table 10-7 cont.

Operating Assumptions

Distribution of revenue by sales elements	Fraction raw revenue	Fraction revenue retained	Fraction collected revenue	Fraction paidout revenue
2nd year				
Retail, direct	0.61	1.00	1.00	0.00
Wholesale, direct	0.14	0.80	0.80	0.00
Retail, channel	0.14	0.90	1.00	0.10
Wholesale, channel	0.11	0.70	0.80	0.10
3rd year				
Retail, direct	0.62	1.00	1.00	0.00
Wholesale, direct	0.13	0.80	0.80	0.00
Retail, channel	0.13	0.90	1.00	0.10
Wholesale, channel	0.12	0.70	0.80	0.10
4th year				
Retail, direct	0.63	1.00	1.00	0.00
Wholesale, direct	0.12	0.80	0.80	0.00
Retail, channel	0.12	0.90	1.00	0.10
Wholesale, channel	0.13	0.70	0.80	0.10
5th year				
Retail, direct	0.64	1.00	1.00	0.00
Wholesale, direct	0.11	0.80	0.80	0.00
Retail, channel	0.11	0.90	1.00	0.10
Wholesale, channel	0.14	0.70	0.80	0.10

Table 10-8 depicts the kind of financial data you can derive from the model. Supporting these numbers is a balance sheet, cash flows, and statement of operations (not shown).

Table 10-8

Summarized Financial Information

	Cumulative five years	
Total revenues	$2,507,887,112	
Expenses		
Building and CO order costs	$48,335,732	1.93%
Facilities administration and maintenance	$628,775,830	25.07%
SGA	$577,032,872	23.01%
Bad debt expense	$62,697,178	2.50%
Depreciation and amortization	$139,871,188	5.58%
Total costs and expenses	$1,456,712,800	58.09%
Income (loss) from operations	$1,051,174,314	41.91%
Other income (expense)		
Interest and other income	$38,192,560	1.52%
Interest expense	($255,789,176)	-10.20%
Total other income (expense)	($217,596,616)	-8.68%
Income before taxes	$833,577,696	33.24%
Income taxes (39%)	$0	0.00%
Net income (loss)	$833,577,696	33.00%

10.3 Simplified Modeling for the Number of Orders Needed to Make Revenue Goals

We alluded previously to a full-fledged model to study in detail all the intricacies of a business plan for a startup next-generation carrier. Although not presenting formal reports in the form of a balance sheet or statement of operations, we want to illustrate the range of parameters that goes into developing such a plan. We will begin by generalizing the modeling. To do

Table 10-9

Example of Service Mix Over Time

	Year 1 MRC price	Year 1 percentage	Year 2 MRC price	Year 2 percentage	Year *n* . . .
Service 1	$3,800	40%	$3,200	30%	
Service 2	$6,500	25%	$6,000	27%	
Service 3	$8,000	10%	$7,700	12%	
Service 4	$12,000	5%	$11,500	7%	
Service 5	$17,000	2%	$16,800	3%	
Service *n* . . .					

so, we'll focus on one aspect: the number of orders needed to reach a certain revenue level.

As output to the marketing modeling activities previously described, the planner should arrive at a matrix like the one shown in Table 10-9, where there could be dozens of different services.

Modeling the entire panoply of services can be daunting, but it's not impossible. One way to simplify the task is to arrive at a proxy environment that has two equivalent services to deal with. One service could be the weighted average MRC of all costs (for example, above $4,000 × 0.3 + 6,500 × 0.2 + and so on), and another service could be 1.5 times the average. For a typical (self-sustaining) data-oriented next-generation carrier, the average proxy service could turn out to be $4,000 MRC and the other proxy service could be $6,000 MRC. These two services can then be used to compute how many orders are needed to achieve a given bottom line. Two other parameters (leading now to a total of four input variables) are price erosion and service mix change after year 1.

Table 10-10 provides an example. In this example, the revenue goals were as follows:

2001 Goal	2002 Goal	2003 Goal	2004 Goal	2005 Goal
$6,632,700	$53,060,813	$168,973,793	$364,186,911	$650,929,776

The pricing of the two services (shown as T1 = Type 1 and T2 = Type 2), the mix in various years, and the erosion in prices are as follows:

	High-end Service (Type 1)	Mid-end Service (Type 2)		
Percentage in 2001%	0.500	0.500	Erosion factor	0.96
Average 2001 revenue	$6,000	$4,000		
Percentage in 2002	0.550	0.450		
Average 2002 revenue	$5,760	$3,840		
Percentage in 2003	0.600	0.400		
Average 2003 revenue	$5,530	$3,686		
Percentage in 2004	0.650	0.350		
Average 2004 revenue	$5,308	$3,539		
Percentage in 2005	0.700	0.300		
Average 2005 revenue	$5,096	$3,397		

Three kinds of outputs are shown in Table 10-10:

1. A calculation of the number of orders needed for each of the two types, assuming no erosion and the same split as year 1 (namely, it is always fifty-fifty)
2. A calculation of the number of orders needed for each of the two types, assuming some erosion and the same split as year 1 (namely, it is always fifty-fifty)
3. A calculation of the number of orders needed for each of the two types, assuming erosion and a split that changes over time at the fraction just described

The number of orders a year for these three cases is shown in Table 10-10.

Notice that with no erosion a cumulative of 13,942 orders would suffice to reach the revenue goal. When considering erosion, more orders are needed, specifically 16,643 in this case. When the mix changes and high-end orders increase, you would expect that you will need fewer orders (15,625 compared with 16,643); however, the price erosion forces the carrier to continue to bring in more orders (15,625 compared to the nonerosion case).

Table 10-10

Number of Orders Needed for Each Category to Reach Revenue Goal

Without Erosion	**2001 Goal**	**2002 Goal**	**2003 Goal**	**2004 Goal**	**2005 Goal**
	$6,632,700	$53,060,813	$168,973,793	$364,186,911	$650,929,776
2001 T1 ramp/month	**7.1**	$6,122,492	$6,122,492	$6,122,492	$6,122,492
2001 T2 ramp/month	**10.6**	$6,122,492	$6,122,492	$6,122,492	$6,122,492
2001 T ramp/month	**17.7**	$12,244,985	$24,489,969	$36,734,954	$48,979,938
2001 total (ramp)	**212.6**				
2002 T1 ramp/month		**43.6**	$37,676,149	$37,676,149	$37,676,149
2002 T2 ramp/month		**65.4**	$37,676,149	$37,676,149	$37,676,149
2002 T ramp/month		**109.0**	$75,352,299	$150,704,597	$226,056,896
2002 total (ramp)		**1,520.8**			
2003 T1 ramp/month			86.9	$75,116,778	$75,116,778
2003 T2 ramp/month			130.4	$75,116,778	$75,116,778
2003 T ramp/month			**217.4**	$150,233,557	$300,467,113
2003 total (ramp)			**4,129.0**		
2004 T1 ramp/month				135.0	$116,636,373
2004 T2 ramp/month				202.5	$116,636,373
2004 T ramp/month				**337.5**	$233,272,747
2004 total (ramp)				**8,178.9**	
2005 T1 ramp/month					192.1
2005 T2 ramp/month					288.2
2005 T ramp/month					**480.3**
2005 total (ramp)					**13,942.5**
Summary # Orders	**212.6**	**1,520.8**	**4,129.0**	**8,178.9**	**13,942.5**

Table 10-10 cont.

Number of Orders Needed for Each Category to Reach Revenue Goal

With Erosion	**2001 Goal**	**2002 Goal**	**2003 Goal**	**2004 Goal**	**2005 Goal**
	$6,632,700	$53,060,813	$168,973,793	$364,186,911	$650,929,776
2001 T1 ramp/month	**7.1**	$5,877,593	$5,642,489	$5,416,789	$5,200,118
2001 T2 ramp/month	**10.6**	$5,877,593	$5,642,489	$5,416,789	$5,200,118
2001 T ramp/month	**17.7**	$11,755,185	$23,040,163	$33,873,742	$44,273,977
2001 total (ramp)	**212.6**				
2002 T1 ramp/month		**46.0**	$36,603,141	$35,139,015	$33,733,455
2002 T2 ramp/month		**69.0**	$36,603,141	$35,139,015	$33,733,455
2002 T ramp/month		**114.9**	$73,206,282	$143,484,312	$210,951,222
2002 total (ramp)		**1,591.6**			
2003 T1 ramp/month			97.9	$74,864,522	$71,869,941
2003 T2 ramp/month			146.9	$74,864,522	$71,869,941
2003 T ramp/month			**244.8**	$149,729,044	$293,468,926
2003 total (ramp)			**4,529.8**		
2004 T1 ramp/month				161.0	$118,165,299
2004 T2 ramp/month				241.5	$118,165,299
2004 T ramp/month				**402.6**	$236,330,599
2004 total (ramp)				**9,360.5**	
2005 T1 ramp/month					242.8
2005 T2 ramp/month					364.1
2005 T ramp/month					**606.9**
2005 total (ramp)					**16,643.3**
(assumes all items are repriced yearly)					
Summary # Orders	**212.6**	**1,591.6**	**4,529.8**	**9,360.5**	**16,643.3**

With Erosion and Split Change	**2001 Goal**	**2002 Goal**	**2003 Goal**	**2004 Goal**	**2005 Goal**
	$6,632,700	$53,060,813	$168,973,793	$364,186,911	$650,929,776
2001 T1 ramp/month	**7.1**	$5,877,593	$5,642,489	$5,416,789	$5,200,118
2001 T2 ramp/month	**10.6**	$5,877,593	$5,642,489	$5,416,789	$5,200,118
2001 T ramp/month	**17.7**	$11,755,185	$23,040,163	$33,873,742	$44,273,977
2001 total (ramp)	**212.6**				
2002 T1 ramp/month		**50.6**	$40,263,455	$38,652,917	$37,106,800
2002 T2 ramp/month		**62.1**	$32,942,827	$31,625,114	$30,360,109
2002 T ramp/month		**112.6**	$73,206,282	$143,484,312	$210,951,222
2002 total (ramp)		**1,564.1**			
2003 T1 ramp/month		117.5	$89,837,426	$86,243,929	
2003 T2 ramp/month		117.5	$59,891,617	$57,495,953	
2003 T ramp/month		**235.0**	$149,729,044	$293,468,926	
2003 total (ramp)		**4,384.7**			
2004 T1 ramp/month				209.3	$153,614,889
2004 T2 ramp/month				169.1	$82,715,710
2004 T ramp/month				**378.4**	$236,330,599
2004 total (ramp)				**8,925.5**	
2005 T1 ramp/month					339.9
2005 T2 ramp/month					218.5
2005 T ramp/month					**558.4**
2005 total (ramp)					**15,625.8**
(assumes all items are repriced yearly)					
(assumes the service % changes, but only for the new batch of customers for that year)					
Summary # Orders	**212.6**	**1,564.1**	**4,384.7**	**8,925.5**	**15,625.8**

10.4 Key Findings for Net-Income-Generating Carriers

We alluded previously to a full-fledged model to help you study all the intricacies of a business plan for a startup next-generation carrier. This section generalizes the modeling and focuses on one aspect: the cost elements faced by a carrier that wants to run a profit-making business. This section imports in a summarized manner SGA and operations expenses from the extensive model of Section 10.1. These two expenses are in the proximity of 18 and 17 percent, respectively, of the total budget. Ultradetailed modeling might show the SGA as 19 percent and the opex as 20 percent, but for the key points that we want to make here, these values will suffice.

The results of this section are of fundamental importance. They make (or break) the business value proposition for next-generation carriers.

A (small) set of key variables is embodied in this analysis. The important ones are as follows:

- The number of high-end broadband customers who can be secured in a single building
- The MRC that can be (is) charged to the customer
- The CPE cost in the building
- The cost of the spur (also known as lateral) into the building
- The metro backbone ring cost and buildings supportable on a ring (see the following list)
- SGA and operations costs

The supportable buildings per ring are assumed to be as follows:

Two 10-Mbps TLS per building	16 buildings
Four 10-Mbps TLS per building	16 buildings
Two 100-Mbps TLS per building	12 buildings
Four 100-Mbps TLS per building	6 buildings
One 1,000-Mbps TLS per building	2 buildings
Two 1,000-Mbps TLS per building	1 building

Three sets of runs are shown in this section:

- **Run 1** Baseline parameters for a typical situation, with CPE costs of $20,000 per building, amortized over 36 months for 10- and 100-Mbps

TLS and $30,000 for GbE TLS. The spur cost is $1,500 a month and the backbone ring cost is $5,000 a month.

- **Run 2** A breakthrough technology where the CPE cost is cut 100 percent to $10,000 and $15,000. Here is the constant hype from vendors: "I can get you equipment that costs less . . . ." Although this certainly helps, the question remains: How much does it help?
- **Run 3** A breakthrough technology that keeps the CPE costs as in Run 1, but doubles the number of buildings supported on the ring. (More likely than not a technology such as dense WDM [DWDM] that allows this would cost more than $20,000 to $30,000 per building; nonetheless, we give the benefit of the doubt to technology suppliers and retain the same CPE cost.)

10.4.1 Run 1: Baseline Parameters for a Typical Situation

Examination of Table 10-11 shows that the sweet spot is four 100-Mbps TLS customers per building. GbE generates more money, but the safe assumption has to be that you cannot find four GbE TLS customers in a building at this time; hence, you cannot secure the sharing of the spur cost among multiple users.

Figures 10-1 through 10-6 represent the information graphically. Examining them will reveal where cost factors lie for next-generation carriers.

Table 10-11

Results of Run 1

Financial Item	Cost
Two 10-Mbps customers in a building on the backbone ring	
Typical total MRC/building (considers one end) ($5,000 MRC per customer, but only $2,500 is generated at the building in question)	$5,000
Allocated ring cost per building/MRC (16 buildings)	$313
Allocated spur cost per building/MRC	$1,500
Allocated CO equipment cost per building/MRC	$354
Allocated CPE cost per building/MRC	$556
Allocated riser costs per building/MRC	$800
Allocated network management costs per building/MRC	$600
Allocated SGA (= 18% of revenue)	$900
Allocated operations costs (= 17% of revenue)	$850
Total costs	$5,872
Net profit (loss)	($872)
Margin %	−17%

Table 10-11 cont.

Results of Run 1

Financial Item	Cost
Four 10-Mbps customers in a building on the backbone ring	
Typical total MRC/building (considers one end) ($5,000 MRC per customer, but only $2,500 is generated at the building in question)	$10,000
Allocated ring cost per building/MRC (16 buildings)	$313
Allocated spur cost per building/MRC	$1,500
Allocated CO equipment cost per building/MRC	$500
Allocated CPE cost per building/MRC	$556
Allocated riser costs per building/MRC	$1,600
Allocated network management costs per building/MRC	$600
Allocated SGA (= 18% of revenue)	$1,800
Allocated operations costs (= 17% of revenue)	$1,700
Total costs	$8,568
Net profit	$1,432
Margin %	14%
Two 100-Mbps customers in a building on the backbone ring	
Typical total MRC/building (considers one end) ($7,000 MRC per customer, but only $3,500 is generated at the building in question)	$7,000
Allocated ring cost per building/MRC (12 buildings)	$417
Allocated spur cost per building/MRC	$1,500
Allocated CO equipment cost per building/MRC	$354
Allocated CPE cost per building/MRC	$556
Allocated riser costs per building/MRC	$800
Allocated network management costs per building/MRC	$600
Allocated SGA (= 18% of revenue)	$1,260
Allocated operations costs (= 17% of revenue)	$1,190
Total costs	$6,676
Net profit	$324
Margin %	5%
Four 100-Mbps customers in a building on the backbone ring	
Typical total MRC/building (considers one end) ($7,000 MRC per customer, but only $3,500 is generated at the building in question)	$14,000
Allocated ring cost per building/MRC (6 buildings)	$833
Allocated spur cost per building/MRC	$1,500
Allocated CO equipment cost per building/MRC	$500
Allocated CPE cost per building/MRC	$556
Allocated riser costs per building/MRC	$1,600

Table 10-11 cont.

Results of Run 1

Financial Item	Cost
Allocated network management costs per building/MRC	$600
Allocated SGA (= 18% of revenue)	$2,520
Allocated operations costs (= 17% of revenue)	$2,380
Total costs	$10,489
Net profit	$3,511
Margin %	25%
One 1,000-Mbps customer in a building on the backbone ring	
Typical total MRC/building (considers one end) ($14,000 MRC per customer, but only $7,000 is generated at the building in question)	$7,000
Allocated ring cost per building/MRC (2 buildings)	$2,500
Allocated spur cost per building/MRC	$1,500
Allocated CO equipment cost per building/MRC	$500
Allocated CPE cost per building/MRC	$833
Allocated riser costs per building/MRC	$400
Allocated network management costs per building/MRC	$600
Allocated SGA (= 18% of revenue)	$1,260
Allocated operations costs (= 17% of revenue)	$1,190
Total costs	$8,783
Net profit (loss)	($1,783)
Margin %	-25%
Two 1,000-Mbps customers in a building on the backbone ring	
Typical total MRC/building (considers one end) ($14,000 MRC per customer, but only $7,000 is generated at the building in question)	$14,000
Allocated ring cost per building/MRC (1 building equivalent or use WDM)	$5,000
Allocated spur cost per building/MRC	$1,500
Allocated CO equipment cost per building/MRC	$707
Allocated CPE cost per building/MRC	$833
Allocated riser costs per building/MRC	$800
Allocated network management costs per building/MRC	$600
Allocated SGA (= 18% of revenue)	$2,520
Allocated operations costs (= 17% of revenue)	$2,380
Total costs	$14,340
Net profit (loss)	($340)
Margin %	−2%

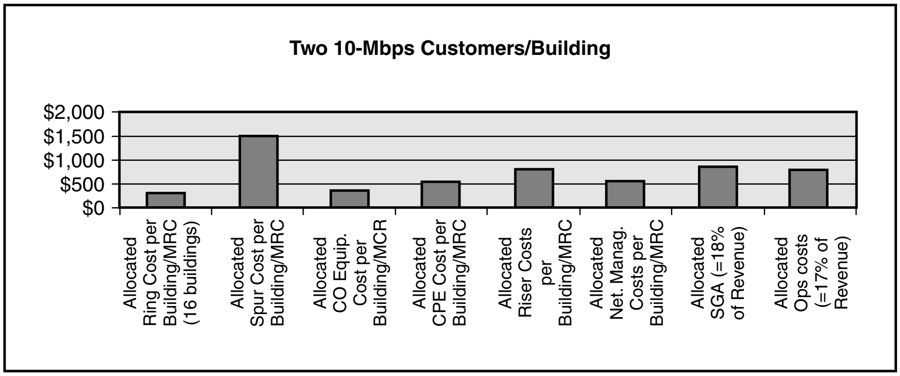


Figure 10-1 Two 10-Mbps customers per building.

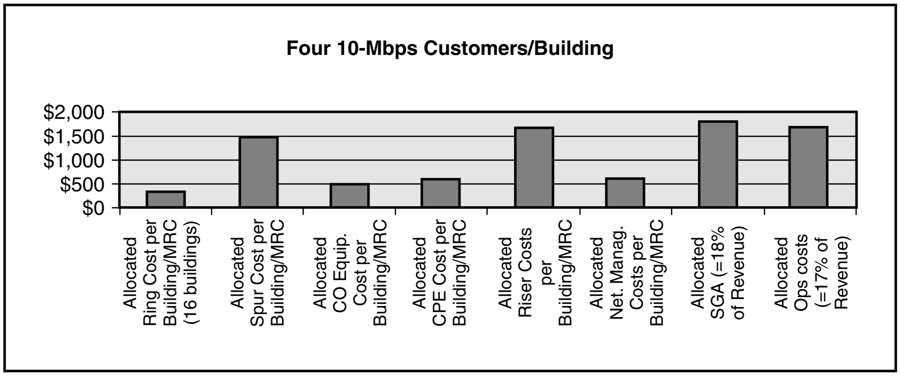


Figure 10-2 Four 10-Mbps customers per building.

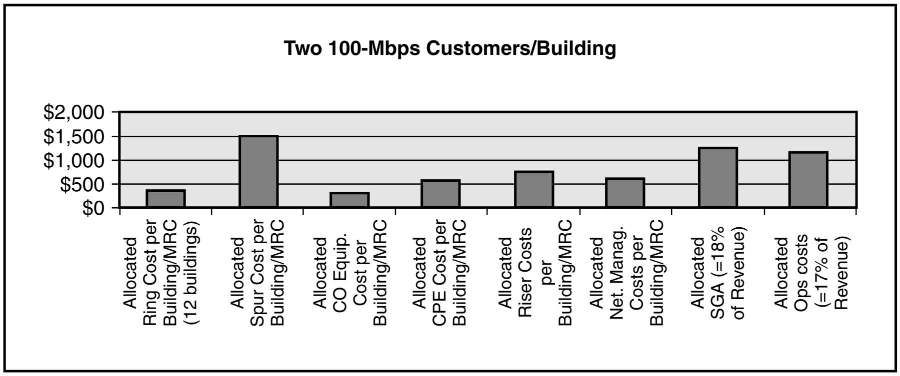


Figure 10-3 Two 100-Mbps customers per building.

Figure 10-4 Four 100-Mbps customers per building.

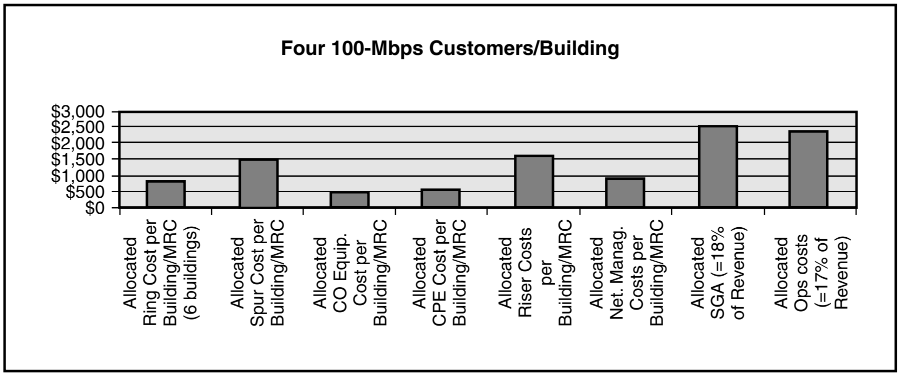


Figure 10-5 One 1,000-Mbps customer per building.

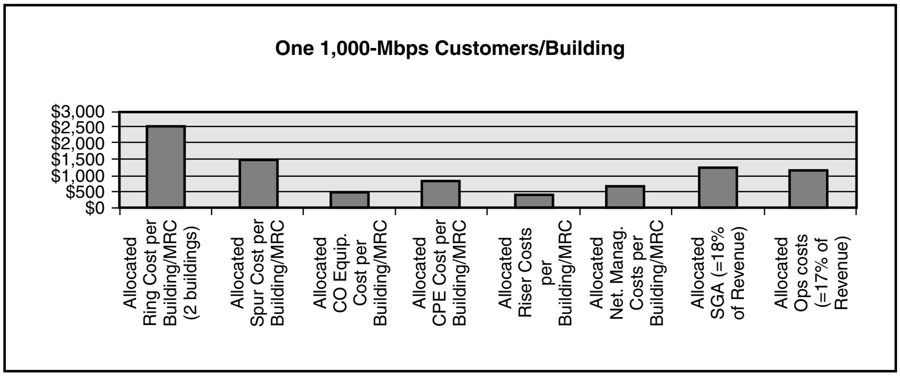
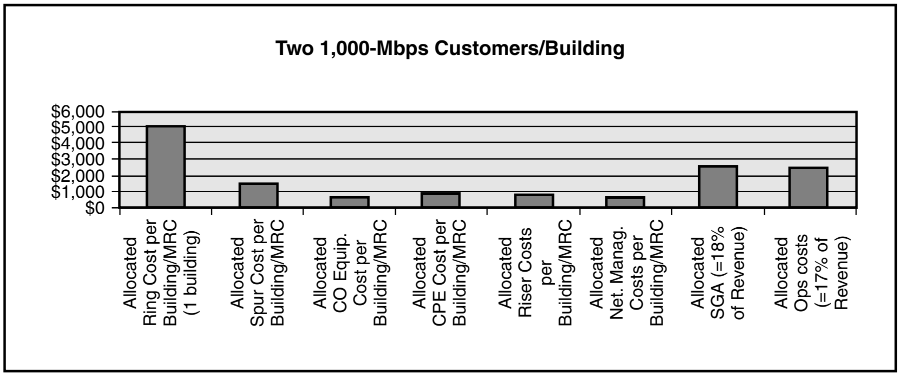


Figure 10-6 Two 1,000-Mbps customers per building.



10.4.2 Run 2: A Breakthrough Technology Where the CPE Cost Is Cut

How important is a breakthrough that keeps the cost per building low? Cost per building is the most critical parameter. Our use of $20,000 and $30,000 for 10- and 100-Mbps TLS and GbE TLS is important because running this model at higher values will show that the proposition quickly becomes untenable from a bottom-line standpoint.

For this run, the same assumptions are used as in the previous case, except that the CPE costs have been reduced by half.

So, how important is it to deprecate (next-generation) SONET in favor of some new technology as journalists in the trade press have a tendency to do? The reader can assess the impact by the tables and figures that follow. The bottom line is an improvement in the profitability of a few percentage points (4 percent); this assumes a 100 percent improvement in the cost of the equipment. See Table 10-12.

Figures 10-7 through 10-12 represent the information in Table 10-12 graphically.

Table 10-12

Results of Run 2

Financial Item	**Cost**
Two 10-Mbps customers in a building on the backbone ring	
Typical total MRC/building (considers one end) ($5,000 MRC per customer, but only $2,500 is generated at the building in question)	$5,000
Allocated ring cost per building/MRC (16 buildings)	$313
Allocated spur cost per building/MRC	$1,500
Allocated CO equipment cost per building/MRC	$354
Allocated CPE cost per building/MRC	$278
Allocated riser costs per building/MRC	$800
Allocated network management costs per building/MRC	$600
Allocated SGA (= 18% of revenue)	$900
Allocated operations costs (= 17% of revenue)	$850
Total costs	$5,594
Net profit (loss)	($594)
Margin %	−12%
Four 10-Mbps customers in a building on the backbone ring	
Typical total MRC/building (considers one end) ($5,000 MRC per customer, but only $2,500 is generated at the building in question)	$10,000

Table 10-12 cont.

Results of Run 2

Financial Item	Cost
Allocated ring cost per building/MRC (16 buildings)	$313
Allocated spur cost per building/MRC	$1,500
Allocated CO equipment cost per building/MRC	$500
Allocated CPE cost per building/MRC	$278
Allocated riser costs per building/MRC	$1,600
Allocated network management costs per building/MRC	$600
Allocated SGA (= 18% of revenue)	$1,800
Allocated operations costs (= 17% of revenue)	$1,700
Total costs	$8,290
Net profit	$1,710
Margin %	17%
Two 100-Mbps customers in a building on the backbone ring	
Typical total MRC/building (considers one end)	$7,000
($7,000 MRC per customer, but only $3,500 is generated at the building in question)	
Allocated ring cost per building/MRC (12 buildings)	$417
Allocated spur cost per building/MRC	$1,500
Allocated CO equipment cost per building/MRC	$354
Allocated CPE cost per building/MRC	$278
Allocated riser costs per building/MRC	$800
Allocated network management costs per building/MRC	$600
Allocated SGA (= 18% of revenue)	$1,260
Allocated operations costs (= 17% of revenue)	$1,190
Total costs	$6,398
Net profit	$602
Margin %	9%
Four 100-Mbps customers in a building on the backbone ring	
Typical total MRC/building (considers one end)	$14,000
($7,000 MRC per customer, but only $3,500 is generated at the building in question)	
Allocated ring cost per building/MRC (6 buildings)	$833
Allocated spur cost per building/MRC	$1,500
Allocated CO equipment cost per building/MRC	$500
Allocated CPE cost per building/MRC	$278
Allocated riser costs per building/MRC	$1,600
Allocated network management costs per building/MRC	$600
Allocated SGA (= 18% of revenue)	$2,520
Allocated operations costs (= 17% of revenue)	$2,380
Total costs	$10,211

Table 10-12 cont.

Results of Run 2

Financial Item	Cost
Net profit	$3,789
Margin %	27%
One 1,000-Mbps customer in a building on the backbone ring	
Typical total MRC/building (considers one end) ($14,000 MRC per customer, but only $7,000 is generated at the building in question)	$7,000
Allocated ring cost per building/MRC (2 buildings)	$2,500
Allocated spur cost per building/MRC	$1,500
Allocated CO equipment cost per building/MRC	$500
Allocated CPE cost per building/MRC	$417
Allocated riser costs per building/MRC	$400
Allocated network management costs per building/MRC	$600
Allocated SGA (= 18% of revenue)	$1,260
Allocated operations costs (= 17% of revenue)	$1,190
Total costs	$8,367
Net profit (loss)	($1,367)
Margin %	−20%
Two 1,000-Mbps customers in a building on the backbone ring	
Typical total MRC/building (considers one end) ($14,000 MRC per customer, but only $7,000 is generated at the building in question)	$14,000
Allocated ring cost per building/MRC (1 building equivalent or use WDM)	$5,000
Allocated spur cost per building/MRC	$1,500
Allocated CO equipment cost per building/MRC	$707
Allocated CPE cost per building/MRC	$417
Allocated riser costs per building/MRC	$800
Allocated network management costs per building/MRC	$600
Allocated SGA (= 18% of revenue)	$2,520
Allocated operations costs (= 17% of revenue)	$2,380
Total costs	$13,924
Net profit	$76
Margin %	1%

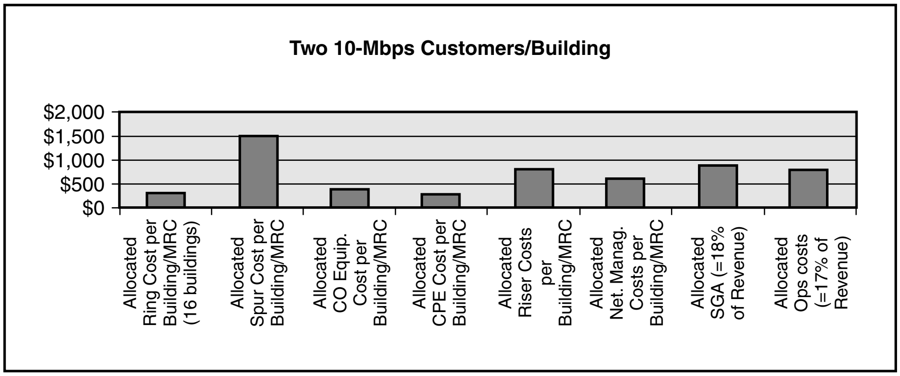


Figure 10-7 Two 10-Mbps customers per building.

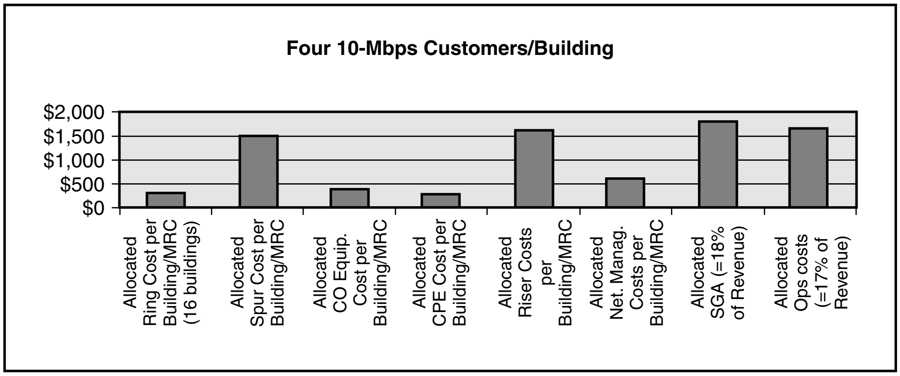


Figure 10-8 Four 10-Mbps customers per building.

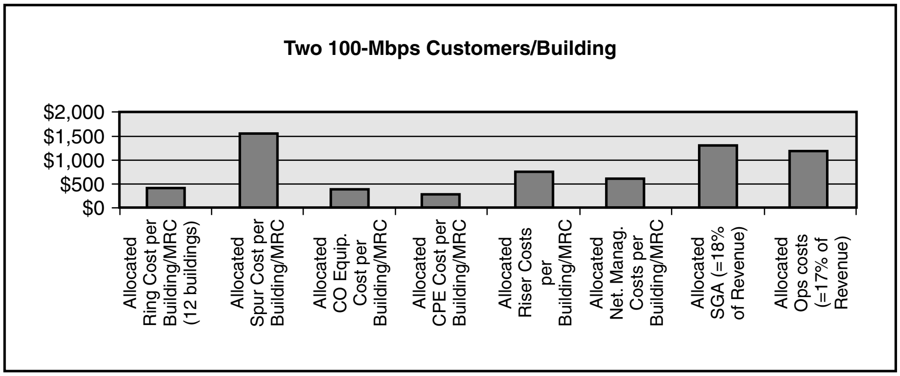


Figure 10-9 Two 100-Mbps customers per building.

Figure 10-10 Four 100-Mbps customers per building.

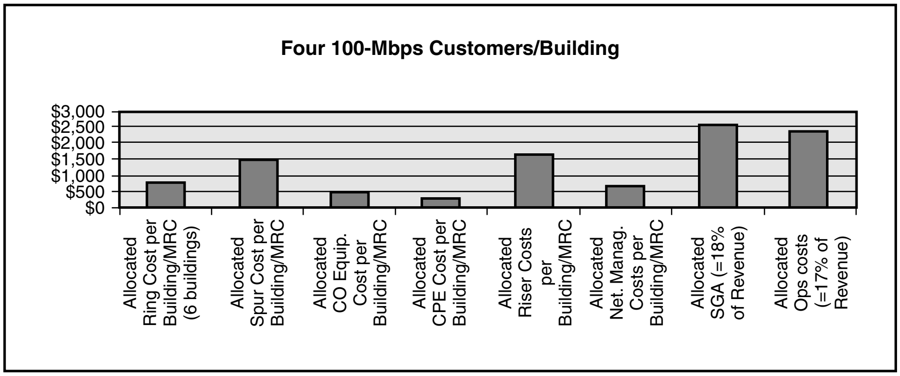


Figure 10-11 One 1,000-Mbps customer per building.

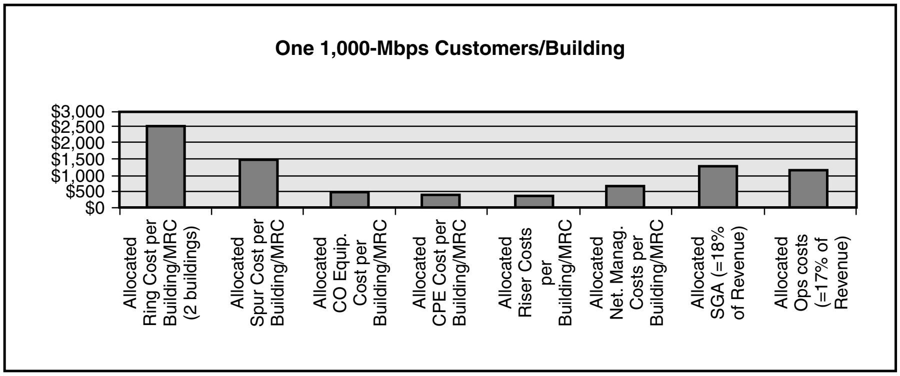
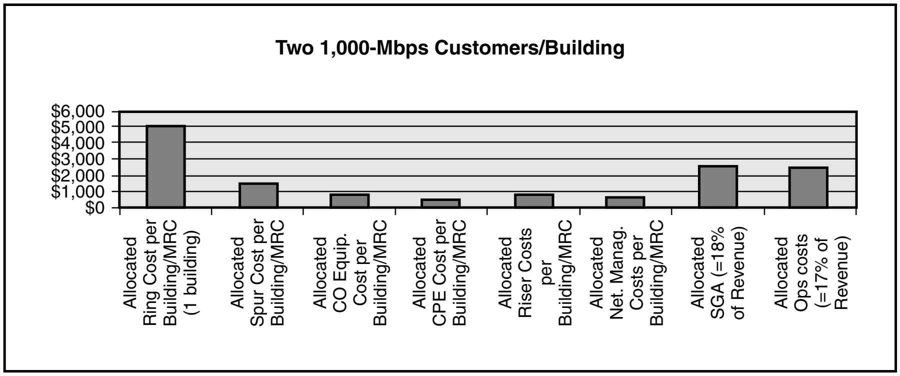


Figure 10-12 Two 1,000-Mbps customers per building.



10.4.3 Run 3: A Breakthrough Technology That Keeps CPE Costs Low While Supporting More Buildings on the Ring

How important is a breakthrough that increases the number of buildings per backbone ring? Typically, you can support 8 to 16 buildings. The same assumptions are used as in the previous case except that the CPE costs have been cut in half.

You can assess the impact by the following tables and figures. The bottom line is an improvement in the profitability of a few single-digit percentage points except at the very high end, namely, if the market existed for a large number of full-throughput GbE services. In this last case, where the backbone cost dominated, the ability to carry twice as many customers improves profitability by about 10 percent. See Table 10-13.

Figures 10-13 through 10-18 represent the information in Table 10-13 graphically.

Table 10-13

Results of Run 3

Financial Item	Cost
Two 10-Mbps customers in a building on the backbone ring	
Typical total MRC/building (considers one end)	$5,000
($5,000 MRC per customer, but only $2,500 is generated at the building in question)	
Allocated ring cost per building/MRC (32 buildings)	$156
Allocated spur cost per building/MRC	$1,500
Allocated CO equipment cost per building/MRC	$354
Allocated CPE cost per building/MRC	$556
Allocated riser costs per building/MRC	$800
Allocated network management costs per building/MRC	$600
Allocated SGA (= 18% of revenue)	$900
Allocated operations costs (= 17% of revenue)	$850
Total costs	$5,715
Net profit (loss)	($715)
Margin %	-14%
Four 10-Mbps customers in a building on the backbone ring	
Typical total MRC/building (considers one end)	$10,000
($5,000 MRC per customer, but only $2,500 is generated at the building in question)	
Allocated ring cost per building/MRC (32 buildings)	$156

Table 10-13

Results of Run 3

Financial Item	Cost
Allocated spur cost per building/MRC	$1,500
Allocated CO equipment cost per building/MRC	$500
Allocated CPE cost per building/MRC	$556
Allocated riser costs per building/MRC	$1,600
Allocated network management costs per building/MRC	$600
Allocated SGA (= 18% of revenue)	$1,800
Allocated operations costs (= 17% of revenue)	$1,700
Total costs	$8,412
Net profit	$1,588
Margin %	16%
Two 100-Mbps customers in a building on the backbone ring	
Typical total MRC/building (considers one end)	$7,000
($7,000 MRC per customer, but only $3,500 is generated at the building in question)	
Allocated ring cost per building/MRC (24 buildings)	$208
Allocated spur cost per building/MRC	$1,500
Allocated CO equipment cost per building/MRC	$354
Allocated CPE cost per building/MRC	$556
Allocated riser costs per building/MRC	$800
Allocated network management costs per building/MRC	$600
Allocated SGA (= 18% of revenue)	$1,260
Allocated operations costs (= 17% of revenue)	$1,190
Total costs	$6,467
Net profit	$533
Margin %	8%
Four 100-Mbps customers in a building on the backbone ring	
Typical total MRC/building (considers one end)	$14,000
($7,000 MRC per customer, but only $3,500 is generated at the building in question)	
Allocated ring cost per building/MRC (12 buildings)	$417
Allocated spur cost per building/MRC	$1,500
Allocated CO equipment cost per building/MRC	$500
Allocated CPE cost per building/MRC	$556
Allocated riser costs per building/MRC	$1,600
Allocated network management costs per building/MRC	$600
Allocated SGA (= 18% of revenue)	$2,520

Table 10-13

Redsults of Run 3

Financial Item	Cost
Allocated operations costs (= 17% of revenue)	$2,380
Total costs	$10,072
Net profit	$3,928
Margin %	28%
One 1,000-Mbps customer in a building on the backbone ring	
Typical total MRC/building (considers one end) ($14,000 MRC per customer, but only $7,000 is generated at the building in question)	$7,000
Allocated ring cost per building/MRC (4 buildings)	$1,250
Allocated spur cost per building/MRC	$1,500
Allocated CO equipment cost per building/MRC	$500
Allocated CPE cost per building/MRC	$833
Allocated riser costs per building/MRC	$400
Allocated network management costs per building/MRC	$600
Allocated SGA (= 18% of revenue)	$1,260
Allocated operations costs (= 17% of revenue)	$1,190
Total costs	$7,533
Net profit (loss)	($533)
Margin %	-8%
Two 1,000-Mbps customers in a building on the backbone ring	
Typical total MRC/building (considers one end) ($14,000 MRC per customer, but only $7,000 is generated at the building in question)	$14,000
Allocated ring cost per building/MRC (2 building equivalent or use WDM)	$2,500
Allocated spur cost per building/MRC	$1,500
Allocated CO equipment cost per building/MRC	$707
Allocated CPE cost per building/MRC	$833
Allocated riser costs per building/MRC	$800
Allocated network management costs per building/MRC	$600
Allocated SGA (= 18% of revenue)	$2,520
Allocated operations costs (= 17% of revenue)	$2,380
Total costs	$11,840
Net profit	$2,160
Margin %	15%

Figure 10-13 Two 10-Mbps customers per building.

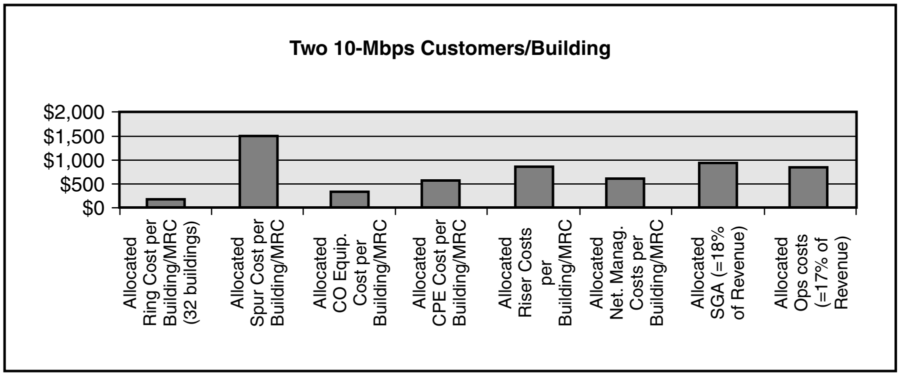


Figure 10-14 Four 10-Mbps customers per building.

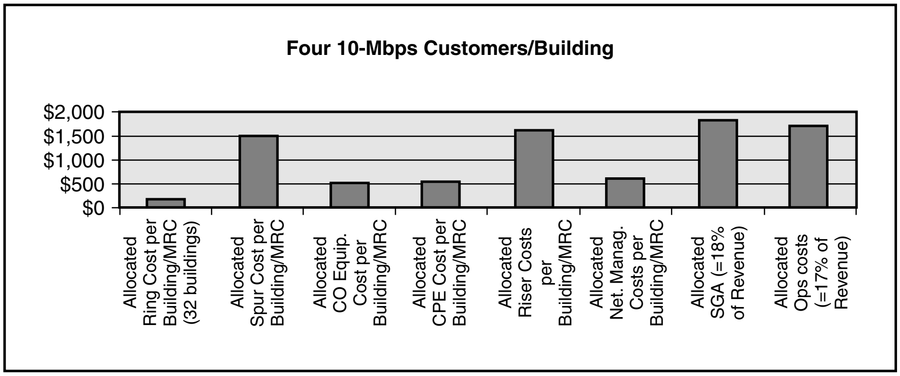


Figure 10-15 Two 100-Mbps customers per building.

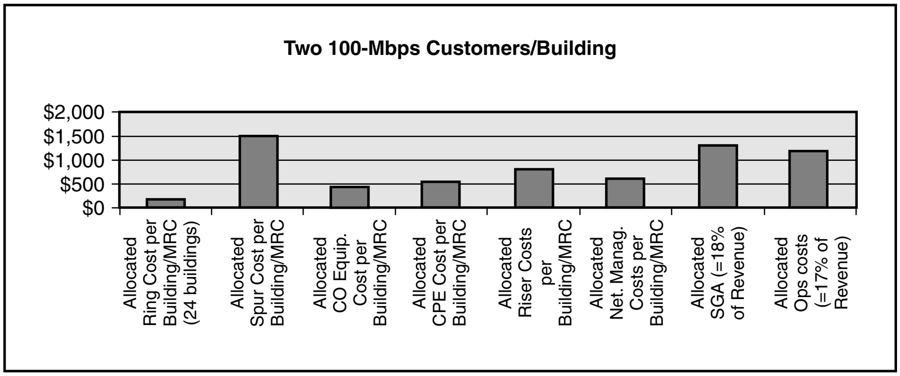


Figure 10-16 Four 100-Mbps customers per building.

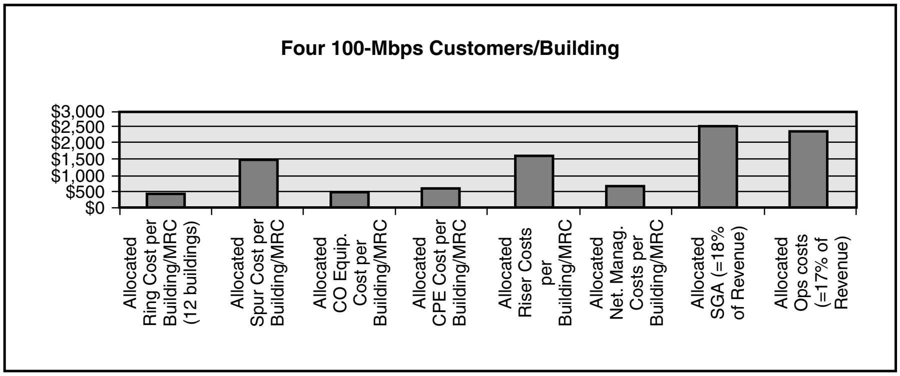


Figure 10-17 One 1,000-Mbps customer per building.

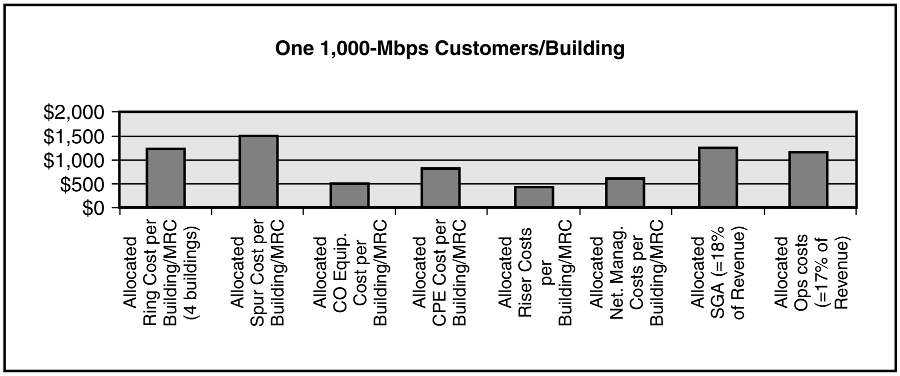
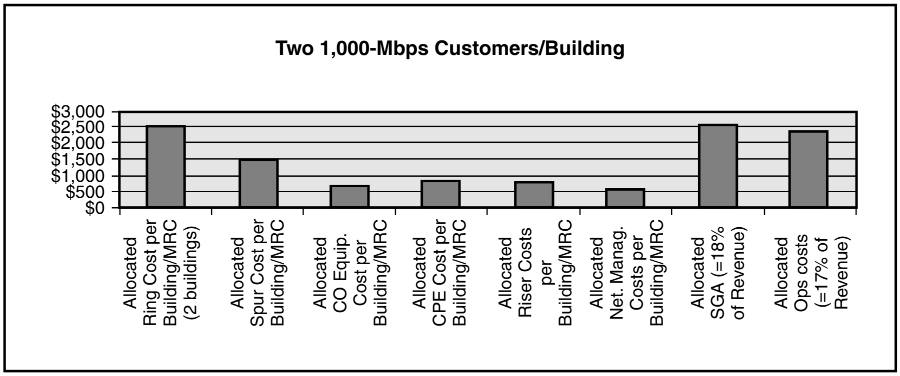


Figure 10-18 Two 1,000-Mbps customers per building.



10.4.4 Key Conclusions

Some major, perhaps critical, points can be synthesized from an analysis of this section. Using the baseline case in Section 10.4.1 and the sweet spot for two or four 100-Mbps TLS customers per building, where a positive profit of 5 to 25 percent can be realized, the cost allocations for a carrier can be determined, as shown in Table 10-14.

Figures 10-19 and 10-20 represent the information in Table 10-14 graphically.

Building access is the dominating cost in a building situation with two customers; SGA and operations costs are the next two costs to follow; fourth in line are the building riser costs (the expense of wiring the building for the two customers); fifth in line are NOC-related costs. Interestingly, none of the hot buttons of the current technology developers address these issues (from this list, a reduction in operations cost would be of interest). None of the bandwidth conservation ideas of Resilient Packet Rings (RPRs) embodied in Switched Multimegabit Data Service (SMDS), IEEE 802.6, and other newer technologies seem to have any relevance to the profit-making line.

SGA and operations costs are the first two dominating costs in a building situation with four customers. In a way, working the building very hard

Table 10-14

Summary of Allocated Costs

Cost Elements	Two Customers per Building	Four Customers per Building
Allocated ring cost per building/MRC	$417	$833
Allocated spur cost per building/MRC	$1,500	$1,500
Allocated CO equipment cost per building/MRC	$354	$500
Allocated CPE cost per building/MRC	$556	$556
Allocated riser costs per building/MRC	$800	$1,600
Allocated network management costs per building/MRC	$600	$600
Allocated SGA (= 18% of revenue)	$1,260	$2,520
Allocated operations costs (= 17% of revenue)	$1,190	$2,380
Total costs	**$6,676**	**$10,489**

Figure 10-19 Two 100-Mbps customers per building.

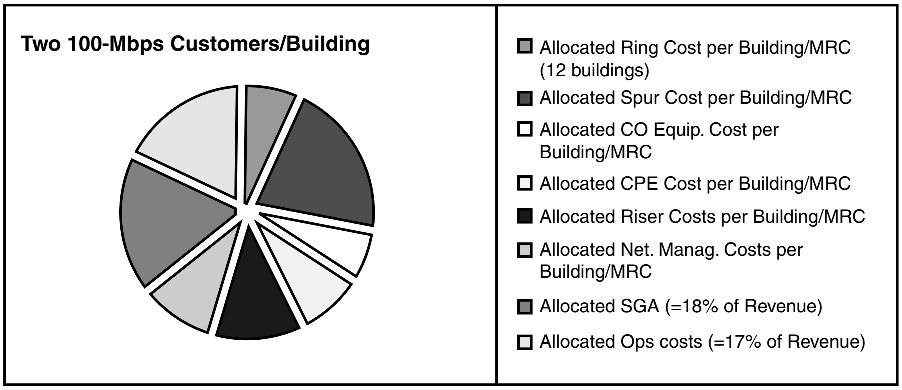
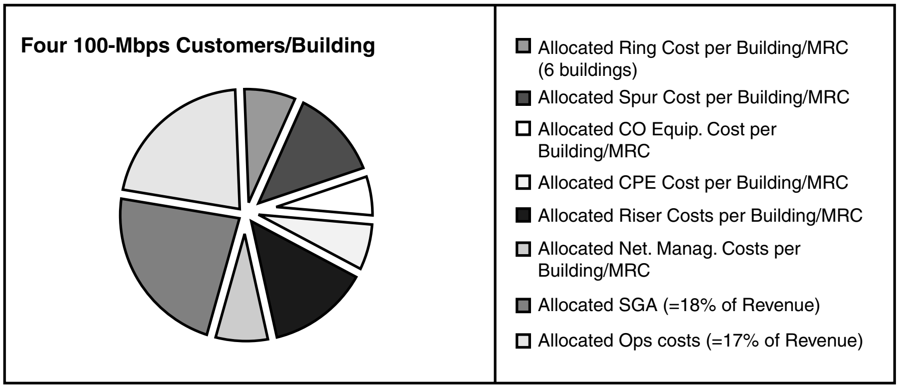


Figure 10-20 Four 100-Mbps customers per building.



to secure four customers might indeed require a lot of marketing, data mining, sales efforts, and so on, raising the cost (in absolute, but not as a fraction). Third in line are building riser costs (the expense of wiring the building for the four customers), which is followed by spur costs in fourth place. Backbone ring costs are in fifth place. Again, interestingly, none of the hot buttons of the current technology developers address the largest four issues.

APPENDIX A

Revenue per Employee

This appendix[1] looks at revenue data for carriers reported for 2001 and computes a weighted average revenue per employee per year. To make the metric more meaningful, we recommend partitioning the carrier space into low end, midrange, and high end based on revenue. This data is collected from public sources, so no guarantees of accuracy are implied. The purpose of the listing is to show the average industry revenue by generic company size, as opposed to results for specific companies.

At startup, a company will inevitably have less revenue and efficiency. Therefore, the targeted revenue per employee per year could start out at $70,000 per year for the first couple of years, grow to $140,000 per year for the next couple of years, and only then move to the target of $290,000. That is to be considered a reasonable growth pattern.

---

[1] Note: This is 2001 data—some companies have been taken over by Chapter 11/7 proceedings.

Table A-1 Revenue per Employee, Looking at Companies with Revenue up to $650 Million

Symbol	Company	Industry	Revenue ($ Million)	Revenue/Employee ($ Thousand)	Market Cap ($ Million)	Cumulative Weighted Average
ACTT	ACT Teleconferencing	Communication equipment	$28	$108	$46	$108,000
ATTL	AT&T Latin America Corp.	Telephone	$40	$44	$473	$70,353
GTTLB	GT Group Telecom Inc.	Computer software	$73	$214	$1,418	$144,723
MFNX	Metromedia Fiber Network	Telephone	$75	$123	$6,197	$137,181
LDIG	Liberty Digital, Inc.	Broadcasting	$84	$201	$1,611	$155,050
SBAC	SBA Communications Corp.	Tele-cell	$87	$149	$1,809	$153,690
INTI	Inet Technologies	Computer software	$159	$347	$1,297	$209,984
AMT	American Tower Corp.	Tele-cell	$258	$147	$6,300	$189,772
ALGX	Allegiance Telecom, Inc.	Telephone	$285	$97	$2,755	$165,493
DCEL	Dobson Communications	Tele-cell	$320	$313	$1,970	$198,994
TWRS	Crown Castle Intl. Corp.	Tele-cell	$346	$276	$5,383	$214,175
MBT	Mobile Telesystems OJSC	Tele-cell	$358	$314	$2,729	$231,088
TSIX	360NETWORKS, INC	Telephone long distance	$386	$386	$9,918	$255,016
IIT	Indonesian Satellite Corp.	Telephone	$386	$226	$1,040	$251,134
MICT	Microcell Telecom.	Tele-cell	$429	$173	$1,639	$241,020
KQIP	KPNQwest N. V.	Telephone	$435	$290	$10,082	$246,703
PTEL	Powertel, Inc.	Tele-cell	$464	$221	$2,263	$243,872
CYCL	Centennial Communications	Tele-cell	$501	$290	$1,850	$248,775
MICC	Millicom Int'l. Cellular	Tele-cell	$552	$103	$1,349	$233,494
PR	Price Communications Corp.	Tele-cell	$552	$373	$1,034	$246,730
						$5,818,000,000

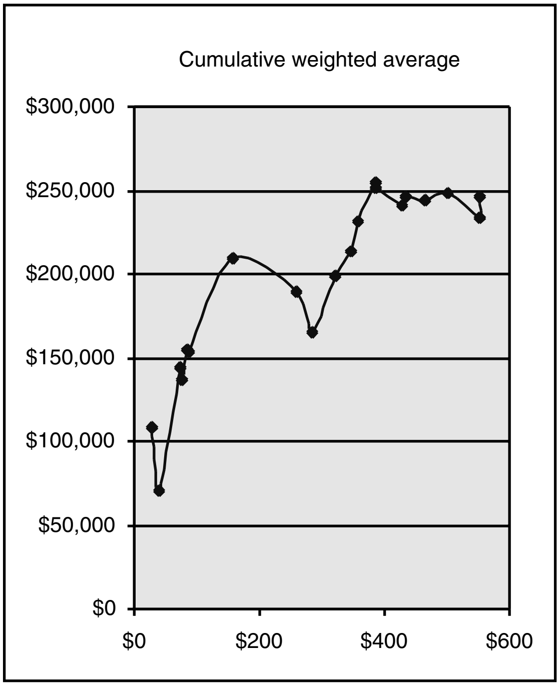


Figure A-1 Cumulative weighted average revenue per employee as a function of company size ($0–$600m).

Table A-2 Revenue per Employee, Looking at Companies with Revenues of $650 Million to $1.3 Billion

Symbol	Company	Industry	Revenue ($ Million)	Revenue/Employee ($ Thousand)	Market Cap ($ Million)	Cumulative Weighted Average
COLT	COLT Telecom Group plc	Telephone	$651	$1,000	$14,911	$1,000,000
ENGSY	Energis PLC	Telephone	$781	$390	$11,345	$667,311
WCG	Williams Communications	Telephone long distance	$839	$91	$7,892	$454,399
BCICF	Bell Canada International	Telephone	$865	$569	$1,396	$486,009
NECSB	NetCom AB	Telephone	$965	$1,107	$4,310	$632,133
TLK	P. T. Telekomunikasi	Telephone	$1,017	$26	$2,908	$511,688
ENT	EQUANT N.V.	Computer software	$1,050	$254	$5,686	$467,821
CZN	Citizens Communications	Telephone	$1,087	$162	$3,933	$422,001
GENU	Genuity Inc.	Computer software	$1,137	$319	$3,348	$408,046
PHI	Philippine Long Distance	Telephone	$1,151	$87	$2,441	$369,324
LVLT	Level 3 Communications	Tele-cell	$1,185	$307	$12,507	$362,439
						$10,728,000,000

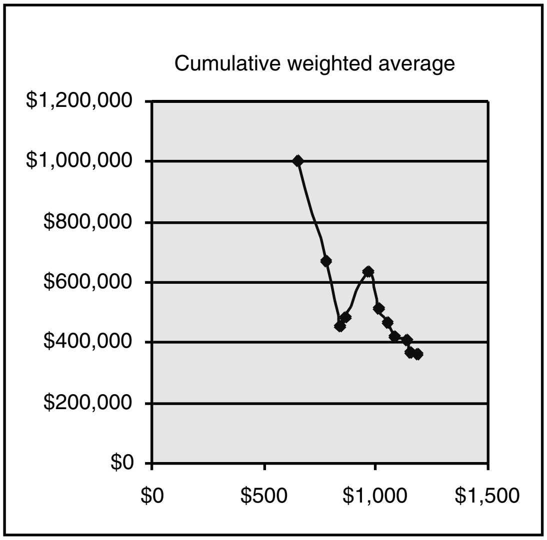


Figure A-2 Cumulative weighted average revenue per employee as a function of company size ($0M–$1.3B).

Table A-3 Revenue per Employee, Looking at Companies with Revenues Between $1.3 and $100 Billion

Symbol	Company	Industry	Revenue ($ Million)	Revenue/Employee ($ Thousand)	Market Cap ($ Million)	Cumulative Weighted Average
MCLD	McLeodUSA, Inc.	Telephone	$1,397	$304	$9,540	$304,000
ATTC	AT&T Canada	Telephone	$1,505	$2,582	$2,804	$1,485,389
MTA	MATAV, Hungarian Telecom.	Telephone	$1,527	$105	$4,097	$1,009,468
CTC	Compania de Telecommunication	Telephone	$1,604	$161	$3,518	$783,885
VNT	Compania Anonima Nacional	Telephone	$1,716	$116	$3,289	$635,983
CTL	Centurytel, Inc.	Telephone	$1,843	$323	$4,093	$575,847
BRW	Broadwing Inc.	Telephone	$2,050	$341	$5,068	$534,493
CHU	China Unicom Limited	Tele-cell	$2,108	$63	$18,824	$462,209
BRP	Brasil Telecom Part. S.A.	Telephone	$3,059	$219	$4,123	$417,949
EMT	Embratel Participacoes	Telephone	$3,060	$298	$4,727	$399,475
NXTL	Nextel Communications	Tele-cell	$3,326	$221	$22,855	$373,883
RG	Rogers Communications	Broadcasting	$3,504	$350	$3,762	$370,749
GX	Global Crossing Ltd.	Telephone long distance	$3,789	$305	$17,252	$362,578
SKM	SK Telecom Co., Ltd.	Tele-cell	$3,910	$542	$13,597	$382,972
CHL	China Mobile Limited	Tele-cell	$4,665	$238	$84,732	$365,660
PCS	Sprint PCS Group	Tele-cell	$6,341	$551	$23,382	$391,544
AT	Alltel Corporation	Telephone	$7,067	$271	$17,227	$375,308
AWE	AT&T Wireless Group	Tele-cell	$7,627	$423	$50,589	$381,361
KTC	Korea Telecom Corp.	Telephone	$10,375	$197	$20,699	$354,219

Symbol	Company	Industry	Revenue ($ Million)	Revenue/Employee ($ Thousand)	Market Cap ($ Million)	Weighted Average
CWP	Cable and Wireless PLC	Telephone	$14,630	$266	$28,876	$339,054
FON	Sprint FON Group	Telephone long distance	$17,688	$227	$21,064	$319,772
BCE	BCE, Inc.	Tele-cell	$18,094	$411	$24,133	$333,427
Q	Qwest Communications Int. (*)	Telephone	$18,954	$189	$66,956	$313,851
BLS	BellSouth Corporation	Telephone	$26,151	$271	$80,215	$307,100
FTE	France Telecom	Telephone	$27,000	$150	$78,076	$285,121
BTY	British Telecommunication	Telephone	$29,757	$217	$57,256	$276,021
DT	Deutsche Telekom AG	Telephone	$36,417	$188	$87,162	$263,652
SBC	SBC Communications Inc.	Telephone	$51,476	$251	$157,638	$261,556
VZ	Verizon Communications	Telephone	$64,707	$248	$141,828	$259,219
T	AT&T Corporation	Telephone long distance	$65,981	$446	$82,950	$287,144
NTT	Nippon Telegraph & Telep.	Telephone	$100,830	$450	$112,723	$317,431
						$542,158,000,000

Figure A-3
Cumulative weighted average revenue per employee as a function of company size ($1.3B–$100B).

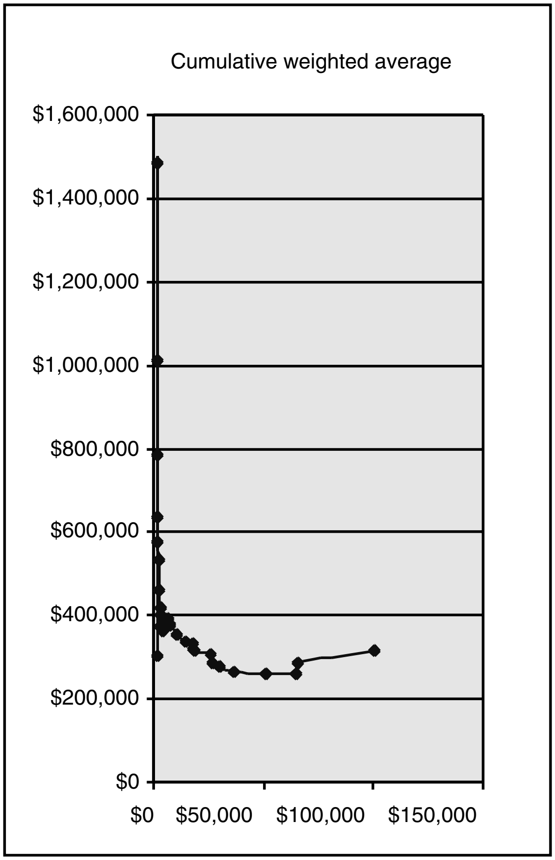


Table A-4 Revenue per Employee, Looking at a Whole Range of Companies

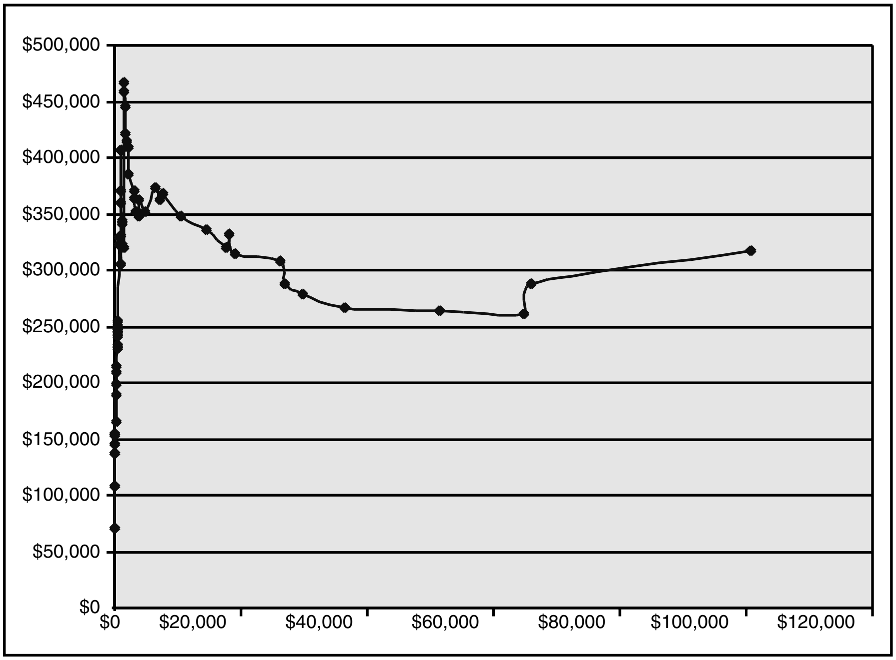
Symbol	Company	Industry	Revenue ($ Million)	Revenue/Employee ($ Thousand)	Market Cap ($ Million)	Cumulative Weighted Average
ACTT	ACT Teleconferencing	Communication equipment	$28	$108	$46	$108,000
ATTL	AT&T Latin America Corp.	Telephone	$40	$44	$473	$70,353
GTTLB	GT Group Telecom Inc.	Computer software	$73	$214	$1,418	$144,723
MFNX	Metromedia Fiber Network	Telephone	$75	$123	$6,197	$137,181
LDIG	Liberty Digital, Inc.	Broadcasting	$84	$201	$1,611	$155,050
SBAC	SBA Communications Corp.	Tele-cell	$87	$149	$1,809	$153,690
INTI	Inet Technologies	Computer software	$159	$347	$1,297	$209,984
AMT	American Tower Corp.	Tele-cell	$258	$147	$6,300	$189,772
ALGX	Allegiance Telecom, Inc.	Telephone	$285	$97	$2,755	$165,493
DCEL	Dobson Communications	Tele-cell	$320	$313	$1,970	$198,994
TWRS	Crown Castle Intl. Corp.	Tele-cell	$346	$276	$5,383	$214,175
MBT	Mobile Telesystems OJSC	Tele-cell	$358	$314	$2,729	$231,088
TSIX	360NETWORKS, INC	Telephone long distance	$386	$386	$9,918	$255,016
IIT	Indonesian Satellite Corp.	Telephone	$386	$226	$1,040	$251,134
MICT	Microcell Telecom.	Tele-cell	$429	$173	$1,639	$241,020
KQIP	KPNQwest N. V.	Telephone	$435	$290	$10,082	$246,703
PTEL	Powertel, Inc.	Tele-cell	$464	$221	$2,263	$243,872
CYCL	Centennial Communications	Tele-cell	$501	$290	$1,850	$248,775
MICC	Millicom Int'l. Cellular	Tele-cell	$552	$103	$1,349	$233,494
PR	Price Communications Corp.	Tele-cell	$552	$373	$1,034	$246,730

Table A-4 Revenue per Employee, Looking at a Whole Range of Companies (continued)

Symbol	Company	Industry	Revenue ($ Million)	Revenue/Employee ($ Thousand)	Market Cap ($ Million)	Cumulative Weighted Average
COLT	COLT Telecom Group PLC	Telephone	$651	$1,000	$14,911	$322,534
ENGSY	Energis PLC	Telephone	$781	$390	$11,345	$329,802
WCG	Williams Communications	Telephone long distance	$839	$91	$7,892	$305,033
BCICF	Bell Canada International	Telephone	$865	$569	$1,396	$330,534
NECSB	NetCom AB	Telephone	$965	$1,107	$4,310	$406,075
TLK	P. T. Telekomunikasi	Telephone	$1,017	$26	$2,908	$370,729
ENT	EQUANT N.V.	Computer software	$1,050	$254	$5,686	$360,504
CZN	Citizens Communications	Telephone	$1,087	$162	$3,933	$343,998
GENU	Genuity Inc.	Computer software	$1,137	$319	$3,348	$341,998
PHI	Philippine Long Distance	Telephone	$1,151	$87	$2,441	$322,891
LVLT	Level 3 Communications	Tele-cell	$1,185	$307	$12,507	$321,753
MCLD	McLeodUSA, Inc.	Telephone	$1,397	$304	$9,540	$320,371
ATTC	AT&T Canada	Telephone	$1,505	$2,582	$2,804	$459,324
MTA	MATAV, Hungarian Telecom.	Telephone	$1,527	$105	$4,097	$466,968
CTC	Compania de Telecommunic.	Telephone	$1,604	$161	$3,518	$445,232
VNT	Compania Anonima Nacional	Telephone	$1,716	$116	$3,289	$421,978
CTL	Centurytel, Inc.	Telephone	$1,843	$323	$4,093	$414,999
BRW	Broadwing Inc.	Telephone	$2,050	$341	$5,068	$409,617
CHU	China Unicom Limited	Tele-cell	$2,108	$63	$18,824	$385,500
BRP	Brasil Telecom Part. S.A.	Telephone	$3,059	$219	$4,123	$370,230
EMT	Embratel Participacoes	Telephone	$3,060	$298	$4,727	$364,160
NXTL	Nextel Communications	Tele-cell	$3,326	$221	$22,855	$352,179

Symbol	Company	Industry	Revenue ($ Million)	Revenue/Employee ($ Thousand)	Market Cap ($ Million)	Cumulative Weighted Average
RG	Rogers Communications	Broadcasting	$3,504	$350	$3,762	$352,002
GX	Global Crossing Ltd.	Telephone long distance	$3,789	$305	$17,252	$348,216
SKM	SK Telecom Co., Ltd.	Tele-cell	$3,910	$542	$13,597	$363,089
CHL	China Mobile Limited	Tele-cell	$4,665	$238	$84,732	$352,595
PCS	Sprint PCS Group	Tele-cell	$6,341	$551	$23,382	$372,904
AT	Alltel Corporation	Telephone	$7,067	$271	$17,227	$362,469
AWE	AT&T Wireless Group	Tele-cell	$7,627	$423	$50,589	$368,493
KTC	Korea Telecom Corp.	Telephone	$10,375	$197	$20,699	$348,046
CWP	Cable and Wireless PLC	Telephone	$14,630	$266	$28,876	$336,238
FON	Sprint FON Group	Telephone long distance	$17,688	$227	$21,064	$320,046
BCE	BCE, Inc.	Tele-cell	$18,094	$411	$24,133	$332,021
Q	Qwest Communications Int.	Telephone	$18,954	$189	$66,956	$314,687
BLS	BellSouth Corporation	Telephone	$26,151	$271	$80,215	$308,428
FTE	France Telecom	Telephone	$27,000	$150	$78,076	$288,014
BTY	British Telecommunication	Telephone	$29,757	$217	$57,256	$279,183
DT	Deutsche Telekom AG	Telephone	$36,417	$188	$87,162	$267,139
SBC	SBC Communications Inc.	Telephone	$51,476	$251	$157,638	$264,600
VZ	Verizon Communications	Telephone	$64,707	$248	$141,828	$261,859
T	AT&T Corporation	Telephone long distance	$65,981	$446	$82,950	$288,394
NTT	Nippon Telegraph & Telep.	Telephone	$100,830	$450	$112,723	$317,559
						$558,704,000,000

Figure A-4 Cumulative weighted average revenue per employee as a function of company size (with whole range of revenues).



APPENDIX B

Calculating Your Sales Staff Number

This appendix posits a hypothetical revenue stream and then calculates the number of salespeople needed to achieve it, using the Rule of 78, assuming a given quota (say, $9,000/month/salesperson).

Suppose, for example, that yearly revenue was $13.1 million for year 1, $80.5 million for year 2, and $210.1 million for year 3. If you were to size the sales force for year 2, you could not simply take $80.5 million and divide by 78×9,000=702,000. (The Rule of 78 says that the $9,000 secured in January contributes to the yearly revenue 12 times, that the $9,000 secured in February contributes 11 times, and so on.) If that simple division were done, it would imply that you needed 115 people. The more exact calculation shown in the following section shows that 60 people in year 2 quarter 1 (Y2Q1), 68 in Y2Q2, 80 in Y2Q3, and 92 in Y2Q4 suffice. This is because there is already an exit run rate that does not have to be resold in year 2. (There could be some erosion of that run rate by some figure such as 5 percent that we did not take into account in the following calculation, but that could easily be modeled.)

Assuming the same revenues ($13.1 million for year 1, $80.5 million for year 2, and $210.1 million for year 3) and a revenue per employee of $70,000 for year 1, $70,000 for year 2, and $140,000 for year 3, the total number of company employees is, respectively, 187, 1,500, and 1,501. The number of salespeople is (as approximated from the following table) 31 for year 1, 80 for year 2, and 128 for year 3. This implies that in year 1 there would be 156 people in operations and management and 31 in sales; in year 2 there would be 1,070 people in operations and management and 80 in sales; and in year 3 there would be 1,373 in operations and management and 128 in sales.

Table B-1 Calculation of the Number of Sales AEs

			First Fiscal Year		
	1st Qtr.	**2nd Qtr.**	**3rd Qtr.**	**4th Qtr.**	**Year Total**
Revenue projections	$360,000	$1,656,000	$3,924,000	$7,164,000	$13,104,000
Cumulative revenue projections	$360,000	$2,016,000	$5,940,000	$13,104,000	$13,104,000
Previous run rate	$0	$180,000	$738,000	$1,593,000	
Old revenue from run rate	$0	$540,000	$2,214,000	$4,779,000	
New revenue to be achieved	$360,000	$1,116,000	$1,710,000	$2,385,000	
Employees=revenue units @$9,000	6.7	20.7	31.7	44.2	
Exit run rate	$180,000	$738,000	$1,593,000	$2,785,500	
			Second Fiscal Year		
	1st Qtr.	**2nd Qtr.**	**3rd Qtr.**	**4th Qtr.**	**Year Total**
Revenue projections	$11,376,000	$16,560,000	$22,716,000	$29,844,000	$80,496,000
Cumulative revenue projections	$24,480,000	$41,040,000	$63,756,000	$93,600,000	$93,600,000
Previous run rate	$2,785,500	$4,295,250	$6,132,375	$8,291,813	
Old revenue from run rate	$8,356,500	$12,885,750	$18,397,125	$24,875,438	
New revenue to be achieved	$3,019,500	$3,674,250	$4,318,875	$4,968,563	
Employees=revenue units @$9,000	55.9	68.0	80.0	92.0	
Exit run rate	$4,295,250	$6,132,375	$8,291,813	$10,776,094	
			Third Fiscal Year		
	1st Qtr.	**2nd Qtr.**	**3rd Qtr.**	**4th Qtr.**	**Year Total**
Revenue projections	$37,944,000	$47,016,000	$57,060,000	$68,076,000	$210,096,000
Cumulative revenue projections	$131,544,000	$178,560,000	$235,620,000	$303,696,000	$303,696,000
Previous run rate	$10,776,094	$13,583,953	$16,716,023	$20,171,988	
Old revenue from run rate	$32,328,281	$40,751,859	$50,148,070	$60,515,965	
New revenue to be achieved	$5,615,719	$6,264,141	$6,911,930	$7,560,035	
Employees=revenue units @$9,000	104.0	116.0	128.0	140.0	
Exit run rate	$13,583,953	$16,716,023	$20,171,988	$23,952,006	

Table B-2 *Detailed Revenue Figures and AE Calculations*

Year 1	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.	Jan.	Feb.	Mar.
In-quarter revenue			$360,000			$1,656,000			$3,924,000			$7,164,000
Exit run rate			$180,000			$738,000			$1,593,000			$2,785,500
Cumulative revenue	$60,000	$180,000	$360,000	$726,000	$1,278,000	$2,016,000	$3,039,000	$4,347,000	$5,940,000	$7,930,500	$9,921,000	$13,104,000
In-quarter sales employees needed	6.7			20.7			31.7			44.2		
April business	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000
May business		$60,000	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000
June business			$60,000	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000
July business				$186,000	$186,000	$186,000	$186,000	$186,000	$186,000	$186,000	$186,000	$186,000
August business					$186,000	$186,000	$186,000	$186,000	$186,000	$186,000	$186,000	$186,000
September business						$186,000	$186,000	$186,000	$186,000	$186,000	$186,000	$186,000
October business							$285,000	$285,000	$285,000	$285,000	$285,000	$285,000
November business								$285,000	$285,000	$285,000	$285,000	$285,000
December business									$285,000	$285,000	$285,000	$285,000
January business										$397,500	$397,500	$397,500
February business											$397,500	$397,500
March business												$397,500
April business												
May business												
June business												
July business												
August business												
September business												
October business												
November business												
December business												
January business												
February business												
March business												
April business												
May business												
June business												
July business												
August business												
September business												

Table B-2 *Detailed Revenue Figures and AE Calculations (continued)*

Year 1	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.	Jan.	Feb.	Mar.
October business												
November business												
December business												
January business												
February business												
March business												
Straight salespeople with management	8			24			36			50		

Note: 31.7 people=285,5000/9,000; 44.2 people=397,500/9,000; and so on.

Year 2	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.	Jan.	Feb.	Mar.
In-quarter revenue			$11,376,000			$16,560,000			$22,716,000			$29,844,000
Exit run rate			$4,295,250			$6,132,375			$8,291,813			$10,776,094
Cumulative revenue	$16,392,750	$20,184,750	$24,480,000	$29,387,625	$34,907,625	$41,040,000	$47,892,188	$55,464,188	$63,756,000	$72,875,906	$82,823,906	$93,600,000
In-quarter sales employees needed	55.9			68.0			80.0			92.0		
April business	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000
May business	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000
June business	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000
July business	$186,000	$186,000	$186,000	$186,000	$186,000	$186,000	$186,000	$186,000	$186,000	$186,000	$186,000	$186,000
August business	$186,000	$186,000	$186,000	$186,000	$186,000	$186,000	$186,000	$186,000	$186,000	$186,000	$186,000	$186,000
September business	$186,000	$186,000	$186,000	$186,000	$186,000	$186,000	$186,000	$186,000	$186,000	$186,000	$186,000	$186,000
October business	$285,000	$285,000	$285,000	$285,000	$285,000	$285,000	$285,000	$285,000	$285,000	$285,000	$285,000	$285,000
November business	$285,000	$285,000	$285,000	$285,000	$285,000	$285,000	$285,000	$285,000	$285,000	$285,000	$285,000	$285,000
December business	$285,000	$285,000	$285,000	$285,000	$285,000	$285,000	$285,000	$285,000	$285,000	$285,000	$285,000	$285,000
January business	$397,500	$397,500	$397,500	$397,500	$397,500	$397,500	$397,500	$397,500	$397,500	$397,500	$397,500	$397,500
February business	$397,500	$397,500	$397,500	$397,500	$397,500	$397,500	$397,500	$397,500	$397,500	$397,500	$397,500	$397,500
March business	$397,500	$397,500	$397,500	$397,500	$397,500	$397,500	$397,500	$397,500	$397,500	$397,500	$397,500	$397,500
April business	$503,250	$503,250	$503,250	$503,250	$503,250	$503,250	$503,250	$503,250	$503,250	$503,250	$503,250	$503,250
May business		$503,250	$503,250	$503,250	$503,250	$503,250	$503,250	$503,250	$503,250	$503,250	$503,250	$503,250
June business			$503,250	$503,250	$503,250	$503,250	$503,250	$503,250	$503,250	$503,250	$503,250	$503,250
July business				$612,375	$612,375	$612,375	$612,375	$612,375	$612,375	$612,375	$612,375	$612,375
August business					$612,375	$612,375	$612,375	$612,375	$612,375	$612,375	$612,375	$612,375
September business						$612,375	$612,375	$612,375	$612,375	$612,375	$612,375	$612,375

Table B-2 *Detailed Revenue Figures and AE Calculations (continued)*

Year 2	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.	Jan.	Feb.	Mar.
October business							$719,813	$719,813	$719,813	$719,813	$719,813	$719,813
November business								$719,813	$719,813	$719,813	$719,813	$719,813
December business									$719,813	$719,813	$719,813	$719,813
January business										$828,094	$828,094	$828,094
February business											$828,094	$828,094
March business												$828,094
April business												
May business												
June business												
July business												
August business												
September business												
October business												
November business												
December business												
January business												
February business												
March business												
Straight sales people with management	63			77			90			104		

Note: 55.9 people=503,250/9,000 and so on.

Year 3	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.	Jan.	Feb.	Mar.
In-quarter revenue			$37,944,000			$47,016,000			$57,060,000			$68,076,000
Exit run rate			$13,583,953			$16,716,023			$20,171,988			$23,952,006
Cumulative revenue	$92,208,047	$104,856,047	$131,544,000	$133,067,977	$148,739,977	$178,560,000	$196,428,012	$215,448,012	$235,620,000	$257,051,994	$279,743,994	$303,696,000
In-quarter sales employees needed	104.0			116.0			128.0			140.0		
April business	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000
May business	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000
June business	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000	$60,000
July business	$186,000	$186,000	$186,000	$186,000	$186,000	$186,000	$186,000	$186,000	$186,000	$186,000	$186,000	$186,000
August business	$186,000	$186,000	$186,000	$186,000	$186,000	$186,000	$186,000	$186,000	$186,000	$186,000	$186,000	$186,000
September business	$186,000	$186,000	$186,000	$186,000	$186,000	$186,000	$186,000	$186,000	$186,000	$186,000	$186,000	$186,000

Table B-2 *Detailed Revenue Figures and AE Calculations (continued)*

Year 3	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.	Jan.	Feb.	Mar.
October business	$285,000	$285,000	$285,000	$285,000	$285,000	$285,000	$285,000	$285,000	$285,000	$285,000	$285,000	$285,000
November business	$285,000	$285,000	$285,000	$285,000	$285,000	$285,000	$285,000	$285,000	$285,000	$285,000	$285,000	$285,000
December business	$285,000	$285,000	$285,000	$285,000	$285,000	$285,000	$285,000	$285,000	$285,000	$285,000	$285,000	$285,000
January business	$397,500	$397,500	$397,500	$397,500	$397,500	$397,500	$397,500	$397,500	$397,500	$397,500	$397,500	$397,500
February business	$397,500	$397,500	$397,500	$397,500	$397,500	$397,500	$397,500	$397,500	$397,500	$397,500	$397,500	$397,500
March business	$397,500	$397,500	$397,500	$397,500	$397,500	$397,500	$397,500	$397,500	$397,500	$397,500	$397,500	$397,500
April business	$503,250	$503,250	$503,250	$503,250	$503,250	$503,250	$503,250	$503,250	$503,250	$503,250	$503,250	$503,250
May business	$503,250	$503,250	$503,250	$503,250	$503,250	$503,250	$503,250	$503,250	$503,250	$503,250	$503,250	$503,250
June business	$503,250	$503,250	$503,250	$503,250	$503,250	$503,250	$503,250	$503,250	$503,250	$503,250	$503,250	$503,250
July business	$612,375	$612,375	$612,375	$612,375	$612,375	$612,375	$612,375	$612,375	$612,375	$612,375	$612,375	$612,375
August business	$612,375	$612,375	$612,375	$612,375	$612,375	$612,375	$612,375	$612,375	$612,375	$612,375	$612,375	$612,375
September business	$612,375	$612,375	$612,375	$612,375	$612,375	$612,375	$612,375	$612,375	$612,375	$612,375	$612,375	$612,375
October business	$719,813	$719,813	$719,813	$719,813	$719,813	$719,813	$719,813	$719,813	$719,813	$719,813	$719,813	$719,813
November business	$719,813	$719,813	$719,813	$719,813	$719,813	$719,813	$719,813	$719,813	$719,813	$719,813	$719,813	$719,813
December business	$719,813	$719,813	$719,813	$719,813	$719,813	$719,813	$719,813	$719,813	$719,813	$719,813	$719,813	$719,813
January business	$828,094	$828,094	$828,094	$828,094	$828,094	$828,094	$828,094	$828,094	$828,094	$828,094	$828,094	$828,094
February business	$828,094	$828,094	$828,094	$828,094	$828,094	$828,094	$828,094	$828,094	$828,094	$828,094	$828,094	$828,094
March business	$828,094	$828,094	$828,094	$828,094	$828,094	$828,094	$828,094	$828,094	$828,094	$828,094	$828,094	$828,094
April business	$935,953	$935,953	$935,953	$935,953	$935,953	$935,953	$935,953	$935,953	$935,953	$935,953	$935,953	$935,953
May business		$935,953	$935,953	$935,953	$935,953	$935,953	$935,953	$935,953	$935,953	$935,953	$935,953	$935,953
June business			$935,953	$935,953	$935,953	$935,953	$935,953	$935,953	$935,953	$935,953	$935,953	$935,953
July business				$1,044,023	$1,044,023	$1,044,023	$1,044,023	$1,044,023	$1,044,023	$1,044,023	$1,044,023	$1,044,023
August business					$1,044,023	$1,044,023	$1,044,023	$1,044,023	$1,044,023	$1,044,023	$1,044,023	$1,044,023
September business						$1,044,023	$1,044,023	$1,044,023	$1,044,023	$1,044,023	$1,044,023	$1,044,023
October business							$1,151,988	$1,151,988	$1,151,988	$1,151,988	$1,151,988	$1,151,988
November business								$1,151,988	$1,151,988	$1,151,988	$1,151,988	$1,151,988
December business									$1,151,988	$1,151,988	$1,151,988	$1,151,988
January business										$1,260,006	$1,260,006	$1,260,006
February business											$1,260,006	$1,260,006
March business												$1,260,006
Straight salespeople with management	117			130			144		157			

Note: No revenue erosion is modeled in these calculations.

APPENDIX C

Generalized "Rule of 78"

This appendix contains a short generalization of the Rule of 78, based on the paper by D. Minoli, "Inductive Formulae For General Sum Operations", *Mathematics of Computation*, Vol. 34, No. 150, April 1980, pp. 543–545. These results have various applications.

1. Introduction. Inductive formulae for sums of powers of consecutive integers are well known; the left side of Table C-1, based on the *CRC Handbook of Tables for Probability and Statistis*[1], depicts such formulae for powers up to 10. These are usually derived directly from the fundamental theorem of sum calculus, or via the Bernoulli polynomials as

$$\sum_{i=0}^{m-1}(1+i)^n = \frac{1}{n+1}\left[B_{n+1}(m+1) - B_{n+1}\right].$$

Formulae for the case where the increment is not one cannot apparently be found explicitly in the literature: [4]–[10]. Conceptually, these formulae are simple to obtain; however, the algebraic manipulations required tend to be overwhelming. In this note we present the first 10 formulae, as obtained on a computer by formal string manipulations.

2. Approach. Let

$$S^n(1, d, m) = \sum_{i=0}^{q}(1+id)^n$$

where $q = (m-1)/d$ is an integer, $d > 0$. We desire a formal closed-form expression for $S^n(1, d, m)$. Clearly

$$\begin{aligned} S^n(1, d, m) &= 1 + \sum_{k=0}^{n}\binom{n}{k}d^{n-k} + \sum_{k=0}^{n}\binom{n}{k}(2d)^{n-k} + \ldots + \sum_{k=0}^{n}\binom{n}{k}(qd)^{n-k} \\ &= \frac{m-1+d}{d} + \sum_{k=0}^{n-1}\binom{n}{k}\left[d^{n-k} + 2^{n-k}d^{n-k} + \cdots + q^{n-k}d^{n-k}\right] \\ &= \frac{m-1+d}{d} + \sum_{k=0}^{n-1}\binom{n}{k}d^{n-k}\sum_{j=1}^{q}j^{n-k}. \end{aligned}$$

Table C-1 Inductive Formulae for General Sum Operations

$$s^1(1, 1, m) = \frac{m}{2}(m + 1)$$

$$s^2(1, 1, m) = \frac{m}{6}(m + 1)(2m + 1)$$

$$s^3(1, 1, m) = \frac{m^2}{4}(m + 1)^2$$

$$s^4(1, 1, m) = \frac{m}{30}(m + 1)(2m + 1)(3m^2 + 3m - 1)$$

$$s^5(1, 1, m) = \frac{m^2}{12}(m + 1)^2(2m^2 + 2m - 1)$$

$$s^6(1, 1, m) = \frac{m}{42}(m + 1)(2m + 1)(3m^4 + 6m^3 - 3m + 1)$$

$$s^7(1, 1, m) = \frac{m^2}{24}(m + 1)^2(3m^4 + 6m^3 - m^2 - 4m + 2)$$

$$s^8(1, 1, m) = \frac{m}{90}(m + 1)(2m + 1)(5m^6 + 15m^5 + 5m^5 + 5m^4 - 15m^3 - m^2 + 9m - 3)$$

$$s^9(1, 1, m) = \frac{m^2}{20}(m + 1)^2(2m^6 + 6m^5 + m^4 - 8m^3 + m^2 + 6m - 3)$$

$$s^{10}(1, 1, m) = \frac{m}{66}(m + 1)(2m + 1)(3m^8 + 12m^7 + 8m^6 - 18m^5 - 10m^4 + 24m^3 + 2m^2 - 15m + 5)$$

The individual terms in the second expression are indeed the entries of the left side of Table 1. The only remaining task is obtaining a formal expression for the first summation by collecting appropriate terms; this is a rather long and tedious task, particularly for high values of n.

The algebraic manipulations have been carried out by a computer program. CPU time on a dedicated DEC PDP 11/70 was 2 hours; the code consisted of about 400 statements. The results are depicted on the right-hand side of Table C-1. We now have closed-form expressions for summations such as $\sum_j(1 + j\sqrt{A})^2$ or $\sum_j(1 + j\pi)^3$.

3. Related Facts.

Fact 1. Besides brute force computation, the results of Table C-1 may be proved by induction on q, for a fixed d, and n.

The following facts can also be proved.

Fact 2. For all n, d,

$$S^n(1, d, m) \approx \frac{m^{n+1} - 1}{(n + 1)d} + \frac{m^n + 1}{2} + \frac{nd}{12}(m^{n-1} - 1),$$

$$s^1(1, d, m) = \frac{1}{2d}(m - 1 + d)(m + 1)$$

$$s^2(1, d, m) = \frac{1}{6d}[m(m + d)(2m + d) - (d - 1)(d - 2)]$$

$$s^3(1, d, m) = \frac{1}{4d}[(m - 1 + d)(m + 1)(m^2 + dm + 1 - d)]$$

$$s^4(1, d, m) = \frac{m^5 - 1}{5d} + \frac{m^4 + 1}{2} + \frac{d}{3}(m^3 - 1) - \frac{d^3}{30}(m - 1)$$

$$s^5(1, d, m) = \frac{m^6 - 1}{6d} + \frac{m^5 + 1}{2} + \frac{5d}{12}(m^4 - 1) - \frac{d^3}{12}(m^2 - 1)$$

$$s^6(1, d, m) = \frac{m^7 - 1}{7d} + \frac{m^6 + 1}{2} + \frac{d}{2}(m^5 - 1) - \frac{d^3}{6}(m^3 - 1) + \frac{d^5}{42}(m - 1)$$

$$s^7(1, d, m) = \frac{m^8 - 1}{8d} + \frac{m^7 + 1}{2} + \frac{7d}{12}(m^6 - 1) - \frac{7d^3}{24}(m^4 - 1) + \frac{d^5}{12}(m^2 - 1)$$

$$s^8(1, d, m) = \frac{m^9 - 1}{9d} + \frac{m^8 + 1}{2} + \frac{2d}{3}(m^7 - 1) - \frac{7d^3}{15}(m^5 - 1) + \frac{2d^5}{9}(m^3 - 1) - \frac{d^7}{30}(m - 1)$$

$$s^9(1, d, m) = \frac{m^{10} - 1}{10d} + \frac{m^9 + 1}{2} + \frac{3d}{4}(m^8 - 1) - \frac{7d^3}{10}(m^6 - 1) + \frac{d^5}{2}(m^4 - 1) - \frac{3d^7}{20}(m^2 - 1)$$

$$s^{10}(1, d, m) = \frac{m^{11} - 1}{11d} + \frac{m^{10} + 1}{2} + \frac{5d}{6}(m^9 - 1) - d^3(m^7 - 1) + d^5(m^5 - 1) - \frac{d^7}{2}(m^3 - 1) + \frac{5d^9}{66}(m - 1)$$

---

which is exact for $n = 1$.

Fact 3. For sufficiently small d, $f(x)$ Riemann integrable, $m \equiv 1 \bmod d$, and

$$\Lambda^f(1, d, m) = \sum_{i=0}^{(m-1)/d} f(1 + id),$$

there exists a δ such that

$$\left| \Lambda^f(1, d, m) - \frac{1}{d}\int_1^m f(x)dx \right| < \delta.$$

In particular, if $f(x) = x^n$

$$\left| S^n(1, d, m) - \frac{m^{n+1} - 1}{(n + 1)d} \right| < \delta.$$

This is related to the Euler-Maclaurin sum formula[2], and a result on the generalized factorial[3].

Fact 4. The sum of the odd integers up to m is equal to the sum of the cubes of all integers up to m, divided by m^2; namely, $S^3(1, 1, m)/m^2 = S^1(1, 2, m)$.

Endnotes

1. *CRC Handbook of Tables for Probability and Statistics* (W. H. Beyer, Ed.), The Chemical Rubber Co., Cleveland, Ohio, 1966.
2. M. Abramowitz & I. Stegun, *Handbook of Mathematical Functions,* Dover, New York, 1964.
3. D. Minoli, "Asymptotic form for the generalized factorial," *Rev. Colombiana Mt.,* v. 11, 1977.

ACRONYMS

10GbE	10 Gigabit Ethernet
ADM	Add/Drop Multiplexer (or Module)
AIS	Alarm Indication Signal
AIS-L	Alarm Indication Signal-Line
AIS-P	Alarm Indication Signal-Path
ANSI	American National Standards Institute
AON	All Optical Networking
APD	Avalanche Photodetector Diode
APS	Automatic Protection Switching
ASIC	Application-Specific Integrated Circuit
ASK	Amplitude Shift Keying
ASP	Application Service Provider
ATM	Asynchronous Transfer Mode
AWG	Arrayed Waveguide Grating
BER	Bit Error Rate
BGP	Border Gateway Protocol
BIP	Bit Interleave Parity
BIP-8	Bit Interleave Parity-8
BIP-N	Bit Interleave Parity-N
BITS	Building Integrated Timing Source
BLEC	Building Local Exchange Carrier
BLSR	Bidirectional Line-Switched Ring
BWDM	Bidirectional Wavelength Division Multiplexing
CI	Customer Installation
CIG	Common Interest Group
CLEC	Competitive Local Exchange Carrier
CNM	Customer Network Management
CO	Central Office
CoS	Class of Service
CPE	Customer Premises Equipment

CSA	Carrier Serving Area
CWDM	Coarse Wavelength Division Multiplexing
DBA	Dynamic Bandwidth Allocation
DBR	Distributed Bragg Reflector (Laser)
DCF	Dispersion Compensation Fiber
DFB	Distributed Feedback (Laser)
DLEC	Data Local Exchange Carrier (or DSL Local Exchange Carrier)
DLL	Data Link Layer
DOS	Denial of Service
DSF	Dispersion-Shifted Fiber
DSL	Digital Subscriber Line
DSU/CSU	Data Service Unit/Channel Service Unit
DWDM	Dense Wavelength Division Multiplexing
ECDL	External Cavity Diode Laser
EDFA	Erbium-Doped Fiber Amplifier
EFM	Ethernet in the First Mile
EGP	Exterior Gateway Protocol
E-IPS	Enhanced-IPS
ELEC	Ethernet Local Exchange Carrier
EoS	Ethernet over SONET
ERDI-P	Enhanced RDI-Path
ESCON	Enterprise System Connection
EWB	Extended Wavelength Band
FBG	Fiber Bragg Grating
FC	Fibre Channel
FCAPS	Fault-Management, Configuration, Accounting, Performance, and Security
FDDI	Fiber-Distributed Data Interface
FDM	Frequency Division Multiplexing
FE	Fast Ethernet
FEC	Forward Error Correction
FEC	Forwarding Equivalence Class

FERF	Far-End Receive Failure
FICON	Fiber Connection
FOTS	Fiber Optic Transmission System
FP	Fabry-Perot
FSC	Fiber-Switch Capable
FSO	Free-Space Optics
FTTB	Fiber to the Building/Business
FTTC	Fiber to the Cabinet/Curb
FTTH	Fiber to the Home
FTTN	Fiber to the Node
FTTx	Fiber to the x (use one or more other letters in place of x, for example, FTTB)
GbE	Gigabit Ethernet
GFF	Gain Flattening Filter
GFP	Genetic Framing Procedure
GMPLS	Generalized Multiprotocol Label Switching
GRE	Generic Routing Encapsulation
HSGG	Hybrid Sol-Gel Glass
IEEE	Institute of Electrical and Electronics Engineers
IETF	Internet Engineering Task Force
ILEC	Incumbent Local Exchange Carrier
IM	Intensity Modulation
InGaAs	Indium-Gallium-Arsenide
InP	Indium Phosphide
ION	Intelligent Optical Network
IP	Internet Protocol
IPS	Intelligent Protection Switching
ISP	Internet Service Provider
ITU	International Telecommunications Union
IXC	Interexchange Carrier
L2	Layer 2
L2S	Layer 2 Switch
L3	Layer 3

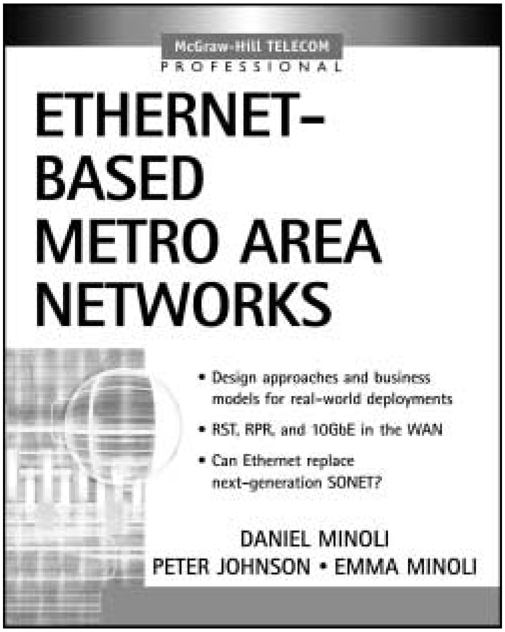
L3S	Layer 3 Switch
LAN	Local Area Network
LD	Laser Diode
LEC	Local Exchange Carrier
LED	Light Emitting Diode
LMP	Link Management Protocol
LOF	Loss of Frame
LOH	Line Overhead
LOL	Loss of Light
LOP-P	Loss of Pointer-Path
LOS	Loss of Signal
LOS-LV	Loss of Signal-Line VT Interface
LSP	Label-Switched Path
LSR	Label-Switched Router
LTE	Line Terminating Equipment
MAC	Media Access Control
MAN	Metropolitan Area Network
MASP	Metropolitan Area Service Provider
MCVD	Modified Chemical Vapor Deposition
MEMS	Micro-Electro Mechanical System
MFD	Mode Field Diameter
MHz	Megahertz
MMF	Multimode Fiber
MOR	Multiwavelength Optical Repeater
MPLS	Multiprotocol Label Switching
MSPP	Multiservice Provisioning Platform
NA	Numerical Aperture
NDSF	Nondispersion-Shifted Fiber
NE	Network Element
NNI	Network Node Interface
NRZ	Nonreturn to Zero
NSP	Network Service Provider
NZDSF	Nonzero Dispersion-Shifted Fiber

OADM	Optical Add/Drop Multiplexer (or Module)
OADN	Optical Add/Drop Node
OAM&P	Operations, Administration, Maintenance, & Provisioning
OC	Optical Carrier
OCVD	Outside Chemical Vapor Deposition
ODSI	Optical Domain Service Interconnect
O-E-O	Optical to Eletrical to Optical
OIF	Optical Internetworking Forum
OLA	Optical Line Amplifier
O-NNI	Optical Network-Node Interface
OSN	Operations Systems Network
OSNR	Optical Signal-to-Noise Ratio
OSPF	Open Shortest Path First
OSS	Operations Support System
OTDM	Optical Time Division Multiplexing
OTN	Optical Transport Network
O-UNI	Optical User-Network Interface
OVPN	Optical Virtual Private Network
OXC	Optical Crossconnect
P2P	Point to Point
PHASIC	Photonic Hybrid Active Silica Integrated Circuit
PHY	Physical Layer
PIN	Positive Intrinsic Negative
PMD	Polarization Mode Dispersion
POH	Path Overhead
PON	Passive Optical Network
POP	Point of Presence
PoS	Packet over SONET
PPP	Point-to-Point Protocol
PRS	Primary Reference Source
PSC	Packet Switch Capable
PTE	Path Terminating Equipment
PTT	Post Telegraph and Telephone

PVC	Permanent Virtual Circuit or Connection
QoS	Quality of Service
RBOC	Regional Bell Operating Company
RDI-L	Remote Defect Indication-Line
RDI-LV	Remote Defect Indication-Line, VT Interface
RDI-P	Remote Defect Indication-Path
REI-L	Remote Error Indication-Line
REI-LV	Remote Error Indication-Line, VT Interface
RIP	Routing Information Protocol
RLEC	Radio Local Exchange Carrier
RPR	Resilient Packet Ring
RRR	Retiming, Reshaping, and Regenerating
RSTA	Rapid Spanning Tree Algorithm
RSTP	Rapid Spanning Tree Protocol
RTS	Roof-Top System
RWA	Routing and Wavelength Assignment
RxM	Receiver Module
RZ	Return to Zero
SAN	Storage Area Network
SBS	Stimulated Brillouin Scattering
SD	Signal Degrade
SDH	Synchronous Digital Hierarchy
SEF	Severely Errored Frame
SEF-LV	Severely Errored Frame-Line, VT Interface
SGA	Sales, General, and Administrative
SLA	Service Level Agreement
SMDS	Switched Multimegabit Data Service
SMF	Single Mode Fiber
SN	Service Node
SNMP	Simple Network Management Protocol
SOH	Section Overhead
SONET	Synchronous Optical Network
SPE	Standard Payload Envelope

Split-MLT	Split-Multilink Trunking Protocol
SPM	Self-Phase Modulation
SRP	Spatial Reuse Protocol
SRS	Stimulated Raman Scattering
SSP	Storage Service Provider
STDM	Statistical Time Division Multiplexing
STE	Section-Terminating Equipment
SVC	Signaled Virtual Circuit or Connection
TDM	Time Division Multiplexing
TE	Traffic Engineering
THz	Terahertz
TIM-P	Trace Identifier Mismatch-Path
TLS	Transparent LAN Service
TOSA	Transmitter Optical Subassembly
TSA	Time Slot Assignment
TSI	Time Slot Interchange
TxM	Transmitter Module
UNI	User-Network Interface
UPSR	Unidirectional Path-Switched Ring
VAD	Vapor Axial Deposition
VC	Virtual Channel
VCSEL	Vertical Cavity Surface Emitting Laser
VLAN	Virtual Local Area Network
VoIP	Voice over Internet Protocol
VPN	Virtual Private Network
VT	Virtual Tributary
WAN	Wide Area Network
WDM	Wavelength Division Multiplexing
WWDM	Wide Wavelength Division Multiplexing
WIS	WAN Interface Sublayer
ZDP	Zero-Dispersion Point

INDEX



Ethernet-Based Metro Area Networks

1st Edition
Daniel Minoli, Peter Johnson, and Emma Minoli
600 pages $69.95 Softcover 0071396861

YOUR SURVIVAL GUIDE TO MAKING SENSE OF THE COMPLEX TECHNOLOGIES AND BURGEONING OPPORTUNITIES OF ETHERNET IN THE METRO ACCESS AND CORE

Suddenly all the action in telecom is on the local segment of the carrier network—the Metropolitan Area Network (MAN). The big challenge in 2002 and beyond will be to rearchitect MANs for broadband services. Of course, great promise comes with great risk and in today's confusing telecom market people are searching for ways to minimize the unknowns.

Ethernet-Based Metro Area Networks is the perfect first resource for assessing your technology options in terms of next-generation platforms, products, and business models. This is the first book to review all emerging metro protocols and their interactions and likely market impact. Written by hands-on industry insiders, *Ethernet-Based Metro Area Networks*

- Sorts out options and reviews each proposed solution against application requirements
- Helps you find your way through complicated regulatory issues
- Offers design approaches and business models for real-world deployments
- Details new and proposed standards for the metro environment
- Critically examines broadband architectures
- Challenges providers to think out of the box

If you've been looking for an informative, plainspoken reference that sheds light on all corners of the metro access and metro core environments—your search ends here.