Lecture Notes
ON
Data Communication and Computer Networks

CoSc2061

Prerequisites: Introduction to Computer Science (CoSc1011)

BSc. II Semester, Regular (CS Second year)

AMBO UNIVERSITY WOLISO CAMPUS, TECHNOLOGY AND INFORMATICS SCHOOL

Department of Computer Science

Prepared By: Abraham A.

Email: abrahamojip210@gmail.com
Phone No.: +251 910272054 or 900272244

2019/20 G.C (2012 E.C)
Woliso, Ethiopia

Course Description and Objectives

This course will explore the various types of the data communication systems, networks and their applications. Concept & terminologies like computer networks, layer architecture (OSI & TCP/IP), network hardware, network software, standardization, network medium, and IP addressing will be explored. The practical aspect will deal with building small to medium level networks including Cabling, Configuring TCP/IP, Peer to Peer Networking, Sharing resources, Client Server Networking.

By the end of this course, you will be able to:

- Understand the concepts and principles of data communications and computer networks
- Understand data transmission and transmission media
- Understand Protocols and various networking components
- Understand TCP/IP & OSI Reference Model
- Understand LAN and WAN technologies
- Understand and implement IP addressing.
- Build small to medium level Computer networks

CHAPTER 1

1. DATA COMMUNICATIONS

When we communicate, we are sharing information. This sharing can be local or remote. Between individuals, local communication usually occurs face to face, while remote communication takes place over distance. The term *telecommunication,* which includes telephony, telegraphy, and television, means communication at a distance *(tele* is Greek for "far").

The word *data* refers to information presented in whatever form is agreed upon by the parties creating and using the data.

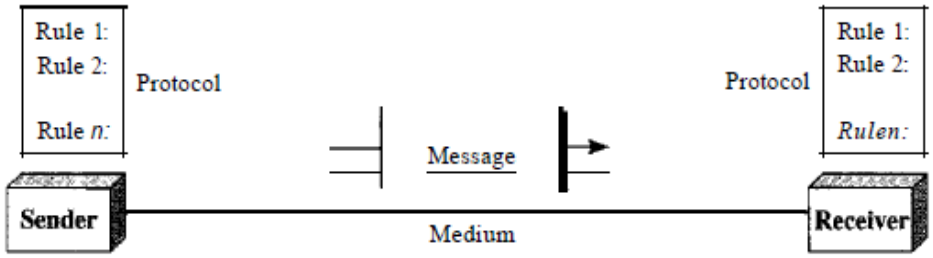
Data communications are the exchange of data between two devices via some form of transmission medium such as a wire cable. For data communications to occur, the communicating devices must be part of a communication system made up of a combination of hardware (physical equipment) and software (programs). The effectiveness of a data communications system depends on four fundamental characteristics: delivery, accuracy, timeliness, and jitter.

1. Delivery. The system must deliver data to the correct destination. Data must be received by the intended device or user and only by that device or user.
2. Accuracy. The system must deliver the data accurately. Data that have been altered in transmission and left uncorrected are unusable.
3. Timeliness. The system must deliver data in a timely manner. Data delivered late are useless. In the case of video and audio, timely delivery means delivering data as they are produced, in the same order that they are produced, and without significant delay. This kind of delivery is called *real-time* transmission.
4. Jitter. Jitter refers to the variation in the packet arrival time. It is the uneven delay in the delivery of audio or video packets. For example, let us assume that video packets are sent every 3D ms. If some of the packets arrive with 3D-ms delay and others with 4D-ms delay, an uneven quality in the video is the result.

Components

A data communications system has five components

1. **Message.** The message is the information (data) to be communicated. Popular forms of information include text, numbers, pictures, audio, and video.
2. **Sender.** The sender is the device that sends the data message. It can be a computer, workstation, telephone handset, video camera, and so on.
3. **Receiver.** The receiver is the device that receives the message. It can be a computer, workstation, telephone handset, television, and so on.
4. **Transmission medium.** The transmission medium is the physical path by which a message travels from sender to receiver. Some examples of transmission media
5. **Protocol.** A protocol is a set of rules that govern data communications. It represents an agreement between the communicating devices. Without a protocol, two devices may be connected but not communicating, just as a person speaking French cannot be understood by a person who speaks only Japanese.



Data Representation Techniques

Information today comes in different forms such as text, numbers, images, audio, and video.

Text

In data communications, text is represented as a bit pattern, a sequence of bits (0's or 1's). Different sets of bit patterns have been designed to represent text symbols. Each set is called a code, and the process of representing symbols is called coding. Today, the prevalent coding system is called Unicode, which uses 32 bits to represent a symbol or character used in any language in the world.

Numbers

Numbers are also represented by bit patterns. However, a code such as ASCII is not used to represent numbers; the number is directly converted to a binary number to simplify mathematical operations.

Images

Images are also represented by bit patterns. In its simplest form, an image is composed of a matrix of pixels (picture elements), where each pixel is a small dot. The size of the pixel depends on the *resolution.* For example, an image can be divided into 1000 pixels or 10,000 pixels. In the second case, there is a better representation of the image (better resolution), but more memory is needed to store the image.

After an image is divided into pixels, each pixel is assigned a bit pattern. The size and the value of the pattern depend on the image. For an image made of only black and- white dots (e.g., a chessboard), a 1-bit pattern is enough to represent a pixel.

Audio

Audio refers to the recording or broadcasting of sound or music. Audio is by nature different from text, numbers, or images. It is continuous, not discrete. Even when we use a microphone to change voice or music to an electric signal, we create a continuous signal. In Chapters 4 and 5, we learn how to change sound or music to a digital or an analog signal.

Video

Video refers to the recording or broadcasting of a picture or movie. Video can either be produced as a continuous entity (e.g., by a TV camera), or it can be a combination of images, each a

discrete entity, arranged to convey the idea of motion. Again we can change video to a digital or an analog signal.

Data Transmission Signals

When data is sent over physical medium, it needs to be first converted into electromagnetic signals. Data itself can be analog such as human voice, or digital such as file on the disk. Both analog and digital data can be represented in digital or analog signals.

Digital Signals

Digital signals are discrete in nature and represent sequence of voltage pulses. Digital signals are used within the circuitry of a computer system.

Analog Signals

Analog signals are in continuous wave form in nature and represented by continuous electromagnetic waves.

Transmission Impairment

When signals travel through the medium, they tend to deteriorate. This may have many reasons as given:

Attenuation

For the receiver to interpret the data accurately, the signal must be sufficiently strong. When the signal passes through the medium, it tends to get weaker. As it covers distance, it loses strength.

Dispersion

As signal travels through the media, it tends to spread and overlaps. The amount of dispersion depends upon the frequency used.

Delay distortion

Signals are sent over media with pre-defined speed and frequency. If the signal speed and frequency do not match, there are possibilities that signal reaches destination in arbitrary fashion. In digital media, this is very critical that some bits reach earlier than the previously sent ones.

Noise

Random disturbance or fluctuation in analog or digital signal is said to be Noise in signal, which may distort the actual information being carried. Noise can be characterized in one of the following class:

Thermal Noise

Heat agitates the electronic conductors of a medium which may introduce noise in the media. Up to a certain level, thermal noise is unavoidable.

Intermodulation

When multiple frequencies share a medium, their interference can cause noise in the medium. Intermodulation noise occurs if two different frequencies are sharing a medium and one of them has excessive strength or the component itself is not functioning properly, then the resultant frequency may not be delivered as expected.

Crosstalk

This sort of noise happens when a foreign signal enters into the media. This is because signal in one medium affects the signal of second medium.

Impulse

This noise is introduced because of irregular disturbances such as lightening, electricity, short-circuit, or faulty components. Digital data is mostly affected by this sort of noise.

Digital Transmission

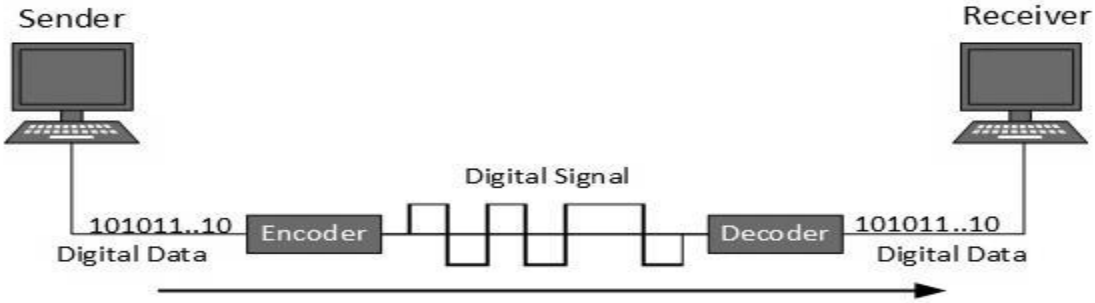
Data or information can be stored in two ways, analog and digital. For a computer to use the data, it must be in discrete digital form. Similar to data, signals can also be in analog and digital form. To transmit data digitally, it needs to be first converted to digital form.

Digital-to-Digital Conversion

This section explains how to convert digital data into digital signals. It can be done in two ways, line coding and block coding. For all communications, line coding is necessary whereas block coding is optional.

Line Coding

The process for converting digital data into digital signal is said to be Line Coding. Digital data is found in binary format. It is represented (stored) internally as series of 1s and 0s.

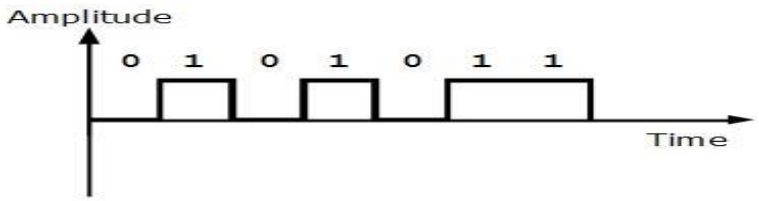


Digital signal is denoted by discreet signal, which represents digital data. There are three types of line coding schemes available:

Line Coding →Uni-Polar
→Polar
→BiPolar

1. **Unipolar Encoding**

Unipolar encoding schemes use single voltage level to represent data. In this case, to represent binary 1, high voltage is transmitted and to represent 0, no voltage is transmitted. It is also called Unipolar-Non-return-to-zero, because there is no rest condition i.e. it either represents 1 or 0.

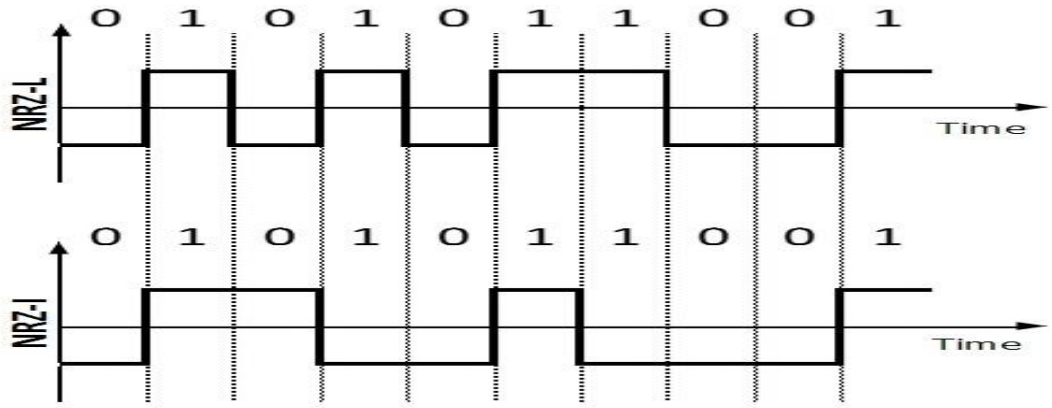


2. **Polar Encoding**

Polar encoding scheme uses multiple voltage levels to represent binary values. Polar encodings is available in four types:

Polar Non Return to Zero (Polar NRZ)

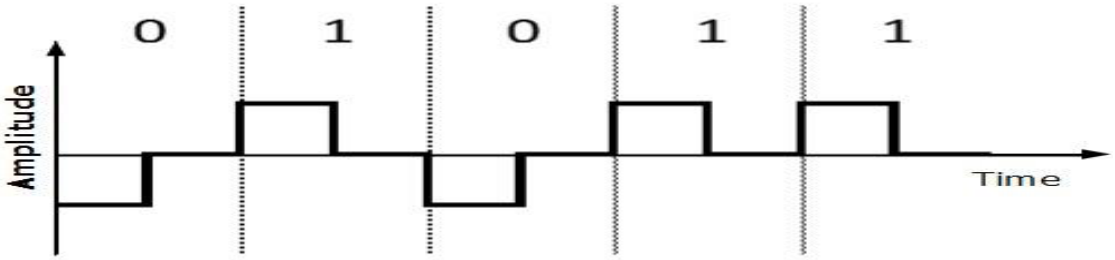
It uses two different voltage levels to represent binary values. Generally, positive voltage represents 1 and negative value represents 0. It is also NRZ because there is no rest condition. NRZ scheme has two variants: NRZ-L and NRZ-I.



NRZ-L changes voltage level at when a different bit is encountered whereas NRZ-I changes voltage when a 1 is encountered.

Return to Zero (RZ)

Problem with NRZ is that the receiver cannot conclude when a bit ended and when the next bit is started, in case when sender and receiver's clock are not synchronized.



RZ uses three voltage levels, positive voltage to represent 1, negative voltage to represent 0 and zero voltage for none. Signals change during bits not between bits.

Manchester

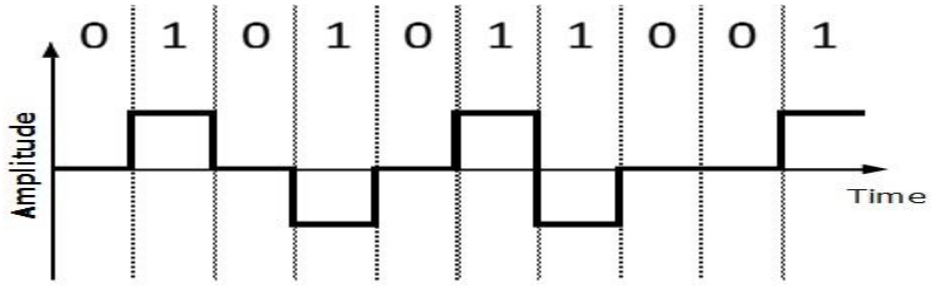
This encoding scheme is a combination of RZ and NRZ-L. Bit time is divided into two halves. It transits in the middle of the bit and changes phase when a different bit is encountered.

Differential Manchester

This encoding scheme is a combination of RZ and NRZ-I. It also transits at the middle of the bit but changes phase only when 1 is encountered.

3. Bipolar Encoding

Bipolar encoding uses three voltage levels, positive, negative, and zero. Zero voltage represents binary 0 and bit 1 is represented by altering positive and negative voltages.



Block Coding

To ensure accuracy of the received data frame, redundant bits are used. For example, in even-parity, one parity bit is added to make the count of 1s in the frame even. This way the original number of bits is increased. It is called Block Coding.

Block coding is represented by slash notation, mB/nB. Means, m-bit block is substituted with n-bit block where n > m. Block coding involves three steps:

1. Division
2. Substitution
3. Combination.

After block coding is done, it is line coded for transmission.

Analog-to-Digital Conversion

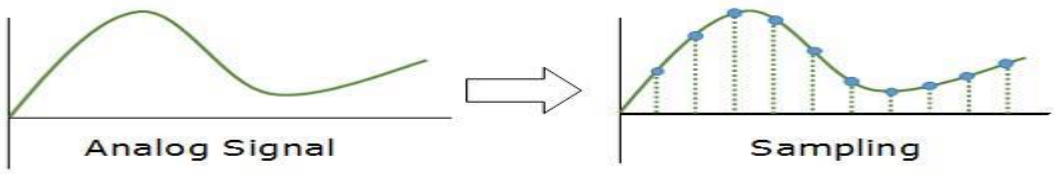
Microphones create analog voice and camera creates analog videos, which are treated is analog data. To transmit this analog data over digital signals, we need analog to digital conversion.

Analog data is a continuous stream of data in the wave form whereas digital data is discrete. To convert analog wave into digital data, we use Pulse Code Modulation (PCM).

PCM is one of the most commonly used method to convert analog data into digital form. It involves three steps:

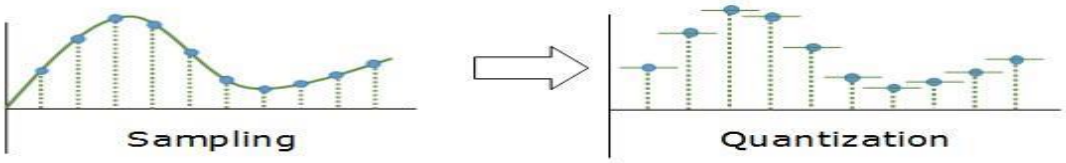
- Sampling
- Quantization
- Encoding.

1. **Sampling**



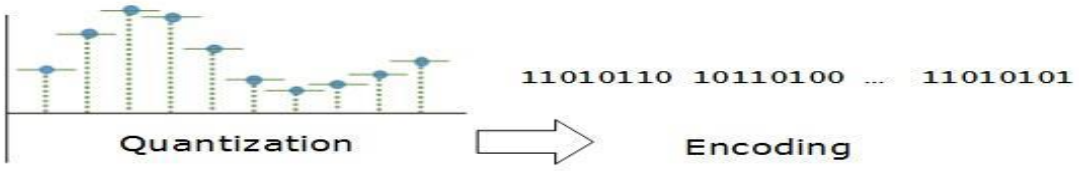
The analog signal is sampled every T interval. Most important factor in sampling is the rate at which analog signal is sampled. According to Nyquist Theorem, the sampling rate must be at least two times of the highest frequency of the signal.

2. **Quantization**



Sampling yields discrete form of continuous analog signal. Every discrete pattern shows the amplitude of the analog signal at that instance. The quantization is done between the maximum amplitude value and the minimum amplitude value. Quantization is approximation of the instantaneous analog value.

3. **Encoding**

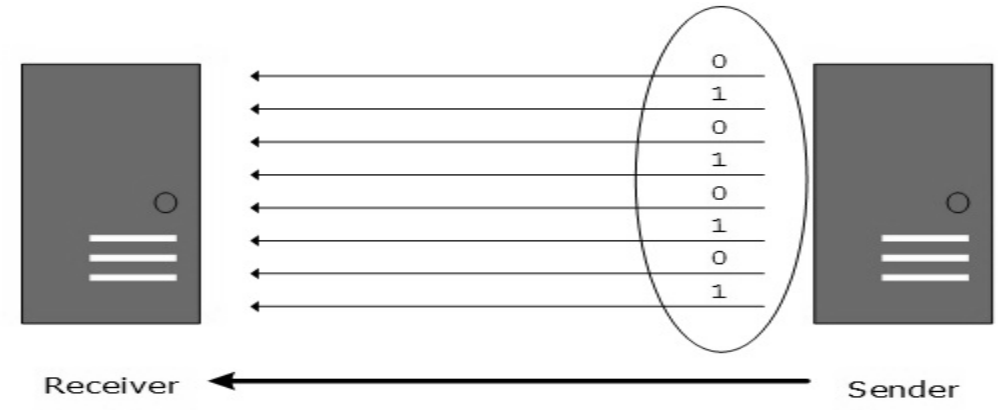


In encoding, each approximated value is then converted into binary format.

Digital Transmission Modes/format

The transmission mode decides how data is transmitted between two computers. The binary data in the form of 1s and 0s can be sent in two different modes: Parallel and Serial.

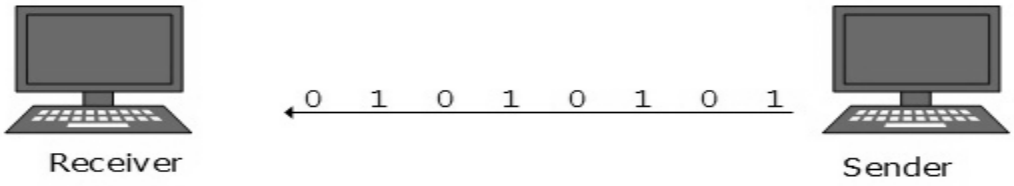
1. **Parallel Transmission**



The binary bits are organized into groups of fixed length. Both sender and receiver are connected in parallel with the equal number of data lines. Both computers distinguish between high order and low order data lines. The sender sends all the bits at once on all lines. Because the data lines are equal to the number of bits in a group or data frame, a complete group of bits (data frame) is sent in one go. Advantage of Parallel transmission is high speed and disadvantage is the cost of wires, as it is equal to the number of bits sent in parallel.

2. **Serial Transmission**

In serial transmission, bits are sent one after another in a queue manner. Serial transmission requires only one communication channel.



Serial transmission can be either asynchronous or synchronous.

- **Asynchronous Serial Transmission**

It is named so because there is no importance of timing. Data-bits have specific pattern and they help receiver recognize the start and end data bits. For example, a 0 is prefixed on every data byte and one or more 1s are added at the end.

Two continuous data-frames (bytes) may have a gap between them.

- **Synchronous Serial Transmission**

Timing in synchronous transmission has importance as there is no mechanism followed to recognize start and end data bits. There is no pattern or prefix/suffix method. Data bits are sent in burst mode without maintaining gap between bytes (8-bits). Single burst of data bits may contain a number of bytes. Therefore, timing becomes very important.
It is up to the receiver to recognize and separate bits into bytes. The advantage of synchronous transmission is high speed, and it has no overhead of extra header and footer bits as in asynchronous transmission.

Analog Transmission

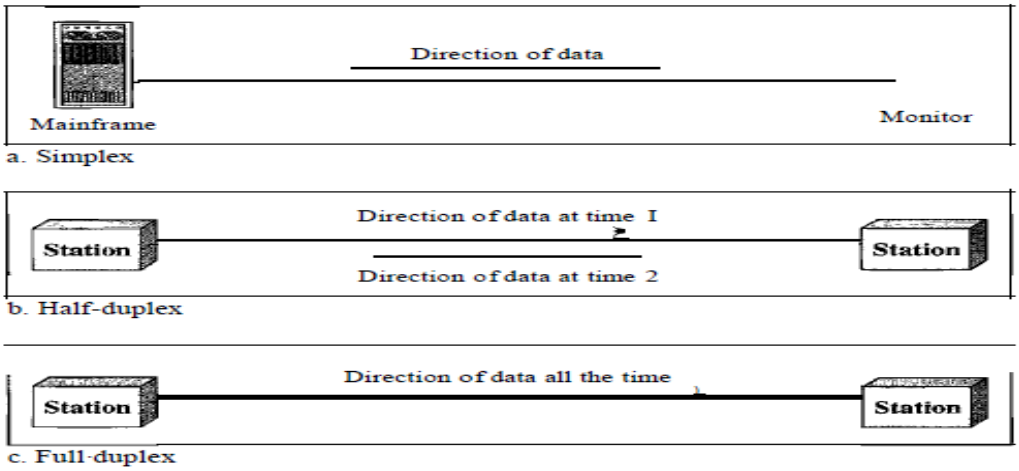
To send the digital data over an analog media, it needs to be converted into analog signal. There can be two cases according to data formatting.

Bandpass: The filters are used to filter and pass frequencies of interest. A bandpass is a band of frequencies which can pass the filter.

Low-pass: Low-pass is a filter that passes low frequencies signals.

When digital data is converted into a bandpass analog signal, it is called digital-to-analog conversion. When low-pass analog signal is converted into bandpass analog signal, it is called analog-to-analog conversion

Modes of Data transmission/Data flow



Shows direction of signal/data flow. Communication between two devices can be simplex, half-duplex, or full-duplex.

Simplex

In simplex mode, the communication is unidirectional, as on a one-way street. Only one of the two devices on a link can transmit; the other can only receive. Keyboards and traditional

monitors are examples of simplex devices. The keyboard can only introduce input; the monitor can only accept output. The simplex mode can use the entire capacity of the channel to send data in one direction. Other example is TV transmission.

Half-Duplex

In half-duplex mode, each station can both transmit and receive, but not at the same time. When one device is sending, the other can only receive, and vice versa. The half-duplex mode is like a one-lane road with traffic allowed in both directions. When cars are traveling in one direction, cars going the other way must wait. In a half-duplex transmission, the entire capacity of a channel is taken over by whichever of the two devices is transmitting at the time. Eg. Military personnel Radio

The half-duplex mode is used in cases where there is no need for communication in both directions at the same time; the entire capacity of the channel can be utilized for each direction.

Full-Duplex

In full-duplex mode (also called duplex), both stations can transmit and receive simultaneously.
The full-duplex mode is like a two-way street with traffic flowing in both directions at the same time. In full-duplex mode, signals going in one direction share the capacity of the link: with signals going in the other direction. This sharing can occur in two ways: Either the link must contain two physically separate transmission paths, one for sending and the other for receiving; or the capacity of the channel is divided between signals traveling in both directions.
One common example of full-duplex communication is the telephone network.
When two people are communicating by a telephone line, both can talk and listen at the same time.
The full-duplex mode is used when communication in both directions is required all the time. The capacity of the channel, however, must be divided between the two directions. E.g Computer network.

Multiplexing

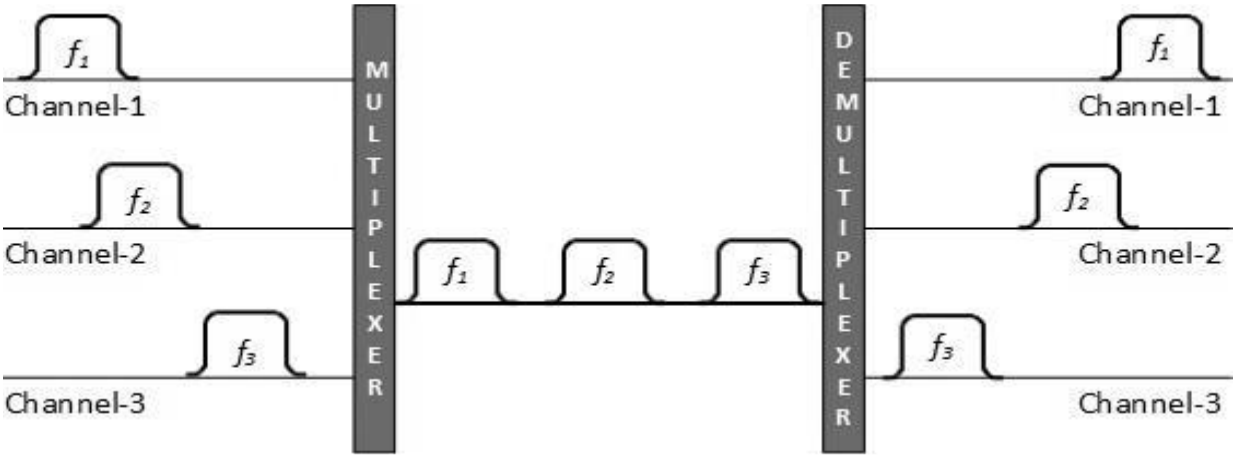
Multiplexing is a technique by which different analog and digital streams of transmission can be simultaneously processed over a shared link. Multiplexing divides the high capacity medium into low capacity logical medium which is then shared by different streams.

Communication is possible over the air (radio frequency), using a physical media (cable), and light (optical fiber). All mediums are capable of multiplexing.

When multiple senders try to send over a single medium, a device called Multiplexer divides the physical channel and allocates one to each. On the other end of communication, a De-multiplexer receives data from a single medium, identifies each, and sends to different receivers.

Frequency Division Multiplexing

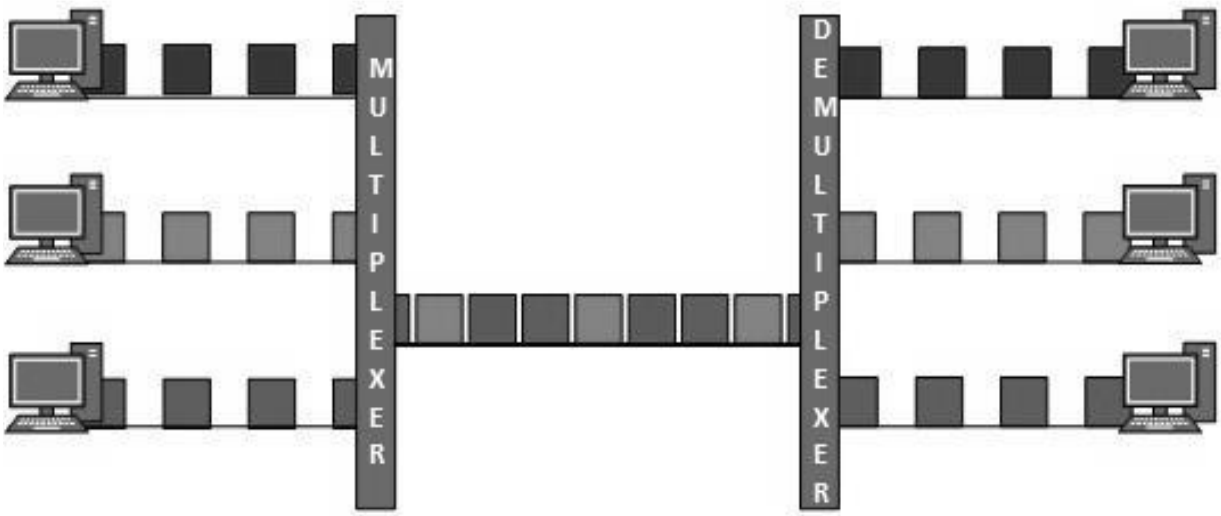
When the carrier is frequency, FDM is used. FDM is an analog technology. FDM divides the spectrum or carrier bandwidth in logical channels and allocates one user to each channel. Each user can use the channel frequency independently and has exclusive access of it. All channels are divided in such a way that they do not overlap with each other. Channels are separated by guard bands. Guard band is a frequency which is not used by either channel.



Time Division Multiplexing

TDM is applied primarily on digital signals but can be applied on analog signals as well. In TDM the shared channel is divided among its user by means of time slot. Each user can transmit data within the provided time slot only. Digital signals are divided in frames, equivalent to time slot i.e. frame of an optimal size which can be transmitted in given time slot.

TDM works in synchronized mode. Both ends, i.e. Multiplexer and De-multiplexer are timely synchronized, and both switch to next channel simultaneously.

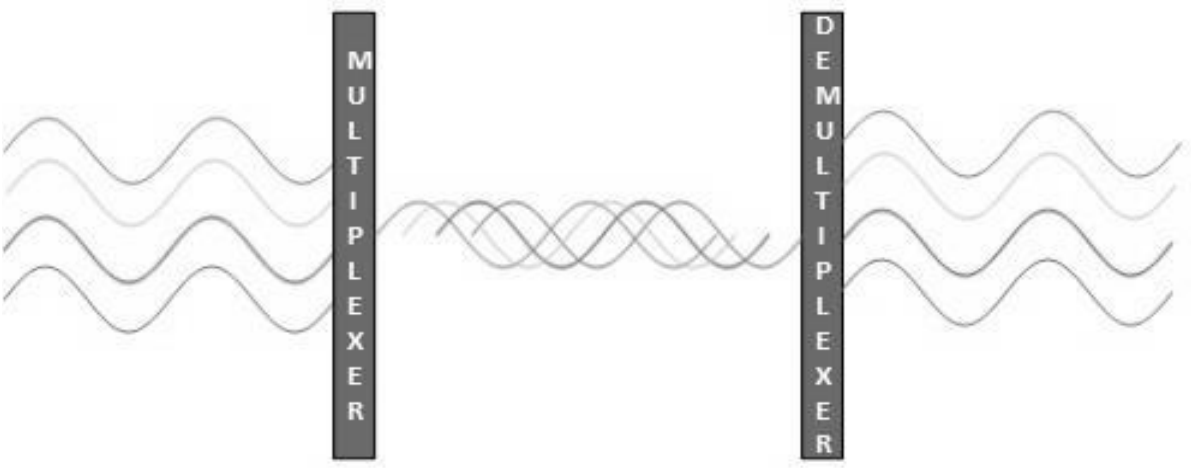


When channel A transmits its frame at one end, the De-multiplexer provides media to channel A on the other end. As soon as the channel A's time slot expires, this side switches to channel B. On the other end, the De-multiplexer works in a synchronized manner and provides media to channel B. Signals from different channels travel the path in interleaved manner.

Wavelength Division Multiplexing

Light has different wavelength (colors). In fiber optic mode, multiple optical carrier signals are multiplexed into an optical fiber by using different wavelengths. This is an analog multiplexing technique and is done conceptually in the same manner as FDM but uses light as signals.

Further, on each wavelength time division multiplexing can be incorporated to accommodate more data signals.



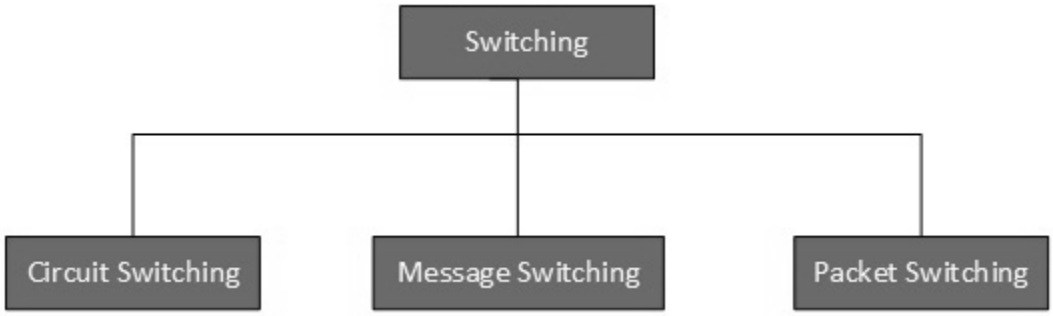
Code Division Multiplexing

Multiple data signals can be transmitted over a single frequency by using Code Division Multiplexing. FDM divides the frequency in smaller channels but CDM allows its users to full bandwidth and transmit signals all the time using a unique code. CDM uses orthogonal codes to spread signals.
Each station is assigned with a unique code, called chip. Signals travel with these codes independently, inside the whole bandwidth. The receiver knows in advance the chip code signal it has to receive.

Switching

Switching is a mechanism by which data/information sent from source towards destination which are not directly connected. Networks have interconnecting devices, which receives data from directly connected sources, stores data, analyze it and then forwards to the next interconnecting device closest to the destination.

Switching can be categorized as:



Data Transmission, Error Detection and Correction

Data-link layer uses some error control mechanism to ensure data bit streams are transmitted with certain level of accuracy. But to understand how errors is controlled, it is essential to know what types of errors may occur.

There may be three types of errors:

- ✓ **Single bit error:** In a frame, there is only one bit, anywhere though, which is corrupt.
- ✓ **Multiple bits error:** Frame is received with more than one bits in corrupted state.
- ✓ **Burst error:** Frame contains more than1 consecutive bits corrupted.

Error control mechanism may involve two possible ways:

- ✓ Error detection
- ✓ Error correction

1. Error Detection :

Errors in the received frames are detected by means of Parity Check and CRC (Cyclic Redundancy Check).

In both scenario, few extra bits are sent along with actual data to confirm that bits received at other end are same as they were sent. If the checks at receiver's end fails, the bits are corrupted.

A. Parity Check

One extra bit is sent along with the original bits to make number of 1s either even. The sender while creating a frame counts the number of 1s in it. **For example,** if even parity is used and number of 1s is even then one bit with value 0 is added. This way number of 1s remain even. Or if the number of 1s is odd, to make it even a bit with value 1 is added.

The receiver simply counts the number of 1s in a frame. If the count of 1s is even and even parity is used, the frame is considered to be not-corrupted and is accepted. If the count of 1s is odd and odd parity is used, the frame is still not corrupted. If a single bit flips in transit, the receiver can detect it by counting the number of 1s. But when more than one bits are in error it is very hard for the receiver to detect the error.

B. ***Cyclic Redundancy Check (CRC)***

Is a different approach to detect if the frame received contains valid data. This technique involves binary division of the data bits being sent. The sender performs a division operation on the bits being sent and calculates the remainder. Before sending the actual bits, the sender adds the remainder at the end of the actual bits. Actual data bits plus the remainder is called a codeword. The sender transmits data bits as codewords.

At the other end, the receiver performs division operation on codewords using the same CRC divisor. If the remainder contains all zeros the data bits are accepted, otherwise there has been some data corruption occurred in transit.

2. **Error Correction**

In digital world, error correction can be done in two ways:

- ✓ **Backward Error Correction:** When the receiver detects an error in the data received, it requests back the sender to retransmit the data unit.
- ✓ **Forward Error Correction**: When the receiver detects some error in the data received, it uses an error-correcting code, which helps it to auto-recover and correct some kinds of errors.

NB: Backward Error Correction, is simple and can only be efficiently used where retransmitting is not expensive, for example fiber optics. But in case of wireless transmission retransmitting may cost too much, in such case Forward Error Correction is used.

CHAPTER TWO
INTRODUCTION TO COMPUTER NETWORKS

Network is a system of interconnected computers and computerized peripherals such as printers is called computer network. This interconnection among computers facilitates information sharing among them. Computers may connect to each other by either wired or wireless media.

Computer network is a computer network is a system in which a number of independent computers are linked together to share data and peripherals, such as files and printers. In the modern world, computer networks have become almost indispensable. All major businesses and governmental and educational institutions make use of computer networks to such an extent that it is now difficult to imagine a world without them.

Computer Network and its Applications

Computer systems and peripherals are connected to form a network. They provide numerous advantages:

- Resource sharing such as printers and storage devices
- Exchange of information by means of e-Mails and FTP
- Information sharing by using Web or Internet
- Interaction with other users using dynamic web pages
- IP phones
- Video conferences
- Parallel computing
- Instant messaging

Classification of Computer Network

Computer networks are classified based on various factors. They include:

- ✓ Geographical span
- ✓ Inter-connectivity
- ✓ Administration
- ✓ Architecture

Geographical Span

Geographically a network can be seen in one of the following categories:

- It may be spanned across your table, among Bluetooth enabled devices, Ranging not more than few meters.
- It may be spanned across a whole building, including intermediate devices to connect all floors.
- It may be spanned across a whole city.
- It may be spanned across multiple cities or provinces.
- It may be one network covering whole world.

Interconnectivity

Components of a network can be connected to each other differently in some fashion. By connectedness we mean either logically, physically, or both ways.

- Every single device can be connected to every other device on network, making the network mesh.
- All devices can be connected to a single medium but geographically disconnected, created bus-like structure.
- Each device is connected to its left and right peers only, creating linear structure.
- All devices connected together with a single device, creating star-like structure.
- All devices connected arbitrarily using all previous ways to connect each other, resulting in a hybrid structure.

Administration

From an administrator's point of view, a network can be private network which belongs a single autonomous system and cannot be accessed outside its physical or logical domain. A network can be public, which is accessed by all.

Network Architecture

Computer networks can be discriminated into various types such as Client-Server, peer-to-peer or hybrid, depending upon its architecture.

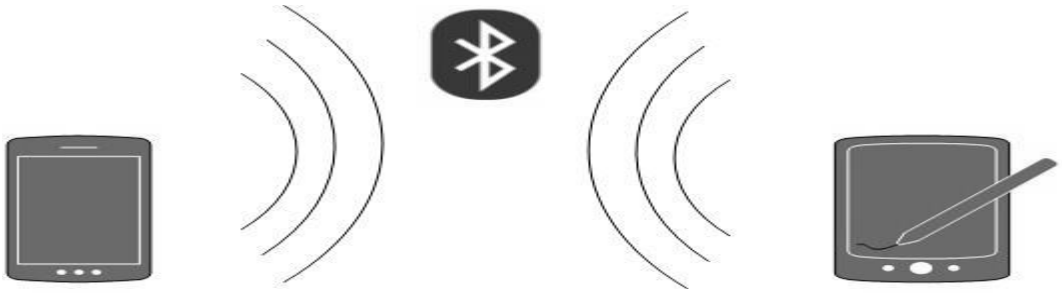
- There can be one or more systems acting as Server. Other being Client, requests the Server to serve requests. Server takes and processes request on behalf of Clients.
- Two systems can be connected Point-to-Point, or in back-to-back fashion. They both reside at the same level and called peers.
- There can be hybrid network which involves network architecture of both the above types.

Types Computer Network

Generally, networks are distinguished based on their geographical span. A network can be as small as distance between your mobile phone and its Bluetooth headphone and as large as the internet itself, covering the whole geographical world.

Personal Area Network

A Personal Area Network (PAN) is smallest network which is very personal to a user. This may include Bluetooth enabled devices or infra-red enabled devices. PAN has connectivity range up to 10 meters. PAN may include wireless computer keyboard and mouse, Bluetooth enabled headphones, wireless printers, and TV remotes.

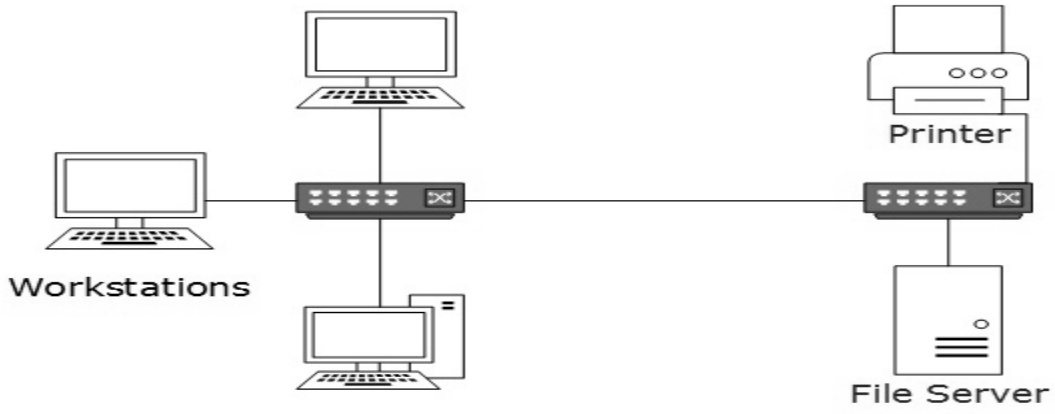


For example, Piconet is Bluetooth-enabled Personal Area Network which may contain up to 8 devices connected together in a master-slave fashion.

Local Area Network

A computer network spanned inside a building and operated under single administrative system is generally termed as Local Area Network (LAN). Usually, LAN covers an organization offices, schools, colleges or universities. Number of systems connected in LAN may vary from as least as two to as much as 16 million.

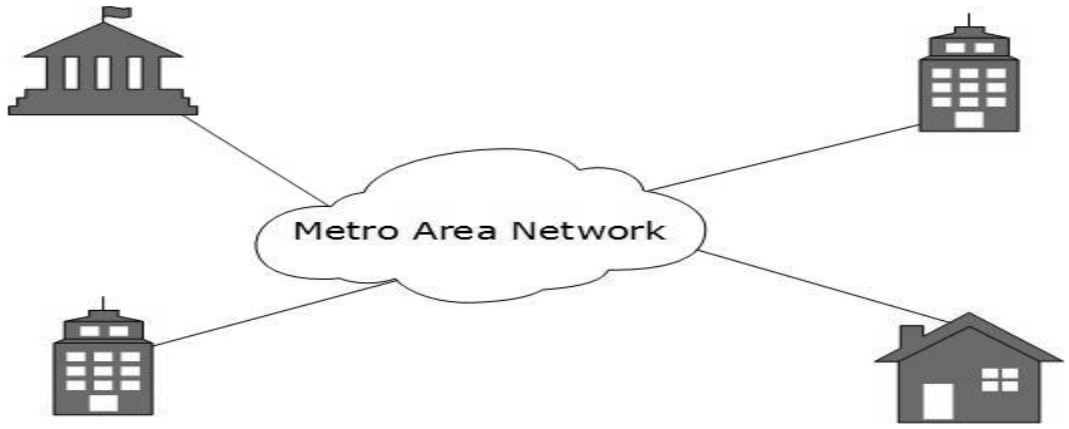
LAN provides a useful way of sharing the resources between end users. The resources such as printers, file servers, scanners, and internet are easily sharable among computers.



LANs are composed of inexpensive networking and routing equipment. It may contains local servers serving file storage and other locally shared applications. It mostly operates on private IP addresses and does not involve heavy routing. LAN works under its own local domain and controlled centrally.
LAN uses either Ethernet or Token-ring technology. Ethernet is most widely employed LAN technology and uses Star topology, while Token-ring is rarely seen.
LAN can be wired, wireless, or in both forms at once.

Metropolitan Area Network
The Metropolitan Area Network (MAN) generally expands throughout a city such as cable TV network. It can be in the form of Ethernet, Token-ring, ATM, or Fiber Distributed Data Interface (FDDI).

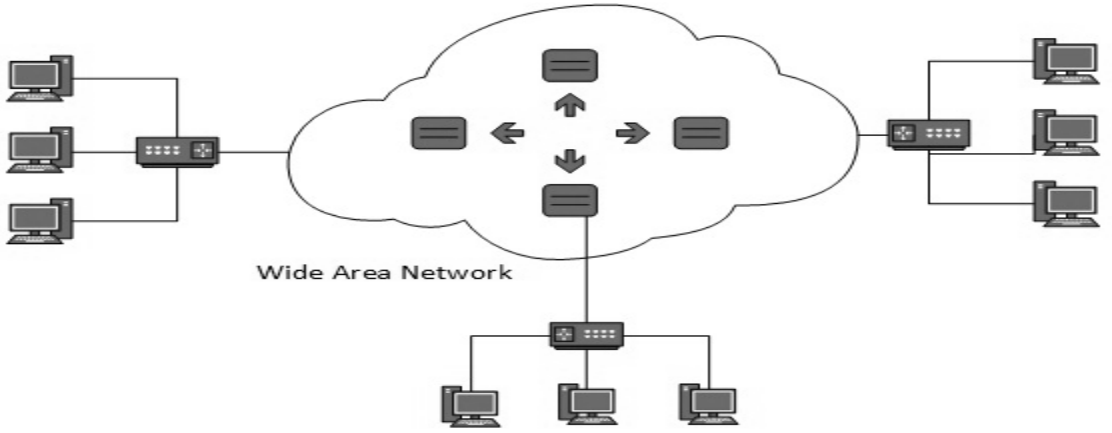
Metro Ethernet is a service which is provided by ISPs. This service enables its users to expand their Local Area Networks. For example, MAN can help an organization to connect all of its offices in a city.



Backbone of MAN is high-capacity and high-speed fiber optics. MAN works in between Local Area Network and Wide Area Network. MAN provides uplink for LANs to WANs or internet.

Wide Area Network

As the name suggests, the Wide Area Network (WAN) covers a wide area which may span across provinces and even a whole country. Generally, telecommunication networks are Wide Area Network. These networks provide connectivity to MANs and LANs. Since they are equipped with very high speed backbone, WANs use very expensive network equipment.



WAN may use advanced technologies such as Asynchronous Transfer Mode (ATM), Frame Relay, and Synchronous Optical Network (SONET). WAN may be managed by multiple administration.

Internetwork

A network of networks is called an internetwork, or simply the internet. It is the largest network in existence on this planet. The internet hugely connects all WANs and it can have connection to LANs and Home networks. Internet uses TCP/IP protocol suite and uses IP as its addressing protocol. Present day, Internet is widely implemented using IPv4. Because of shortage of address spaces, it is gradually migrating from IPv4 to IPv6.
Internet enables its users to share and access enormous amount of information worldwide. It uses WWW, FTP, email services, audio, and video streaming etc. At huge level, internet works on Client-Server model.
Internet uses very high speed backbone of fiber optics. To inter-connect various continents, fibers are laid under sea known to us as submarine communication cable.
Internet is widely deployed on World Wide Web services using HTML linked pages and is accessible by client software known as Web Browsers. When a user requests a page using some web browser located on some Web Server anywhere in the world, the Web Server responds with the proper HTML page. The communication delay is very low.
Internet is serving many proposes and is involved in many aspects of life. Some of them are:

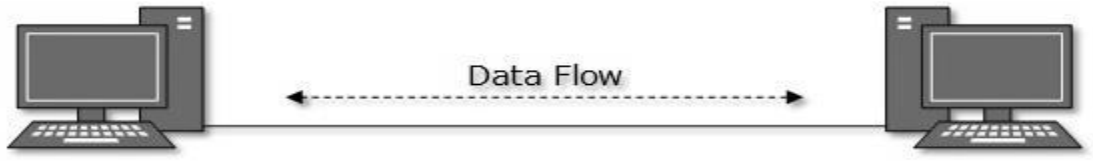
- Web sites
- E-mail
- Instant Messaging
- Blogging
- Social Media
- Marketing
- Networking
- Resource Sharing
- Audio and Video Streaming

Computer Network Topologies

A Network Topology is the arrangement with which computer systems or network devices are connected to each other. Topologies may define both physical and logical aspect of the network. Both logical and physical topologies could be same or different in a same network.

Point-to-Point

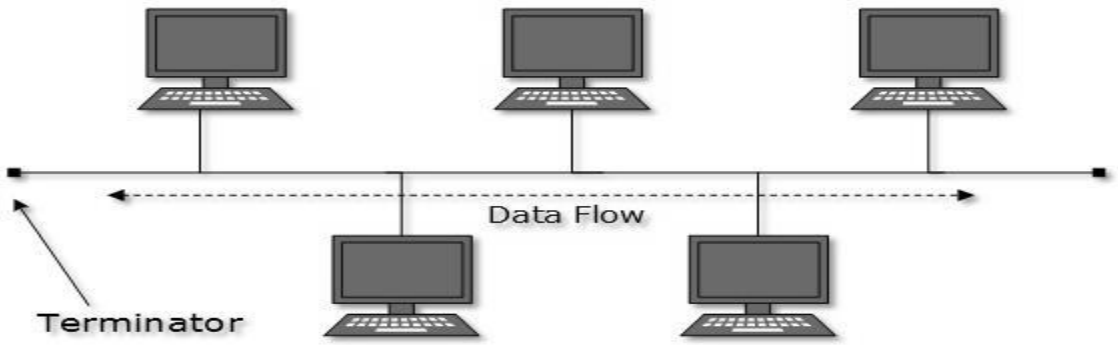
Point-to-point networks contains exactly two hosts such as computer, switches, routers, or servers connected back to back using a single piece of cable. Often, the receiving end of one host is connected to sending end of the other and vice versa.



If the hosts are connected point-to-point logically, then may have multiple intermediate devices. But the end hosts are unaware of underlying network and see each other as if they are connected directly.

Bus Topology

In case of Bus topology, all devices share single communication line or cable. Bus topology may have problem while multiple hosts sending data at the same time. Therefore, Bus topology either uses CSMA/CD technology or recognizes one host as Bus Master to solve the issue. It is one of the simple forms of networking where a failure of a device does not affect the other devices. But failure of the shared communication line can make all other devices stop functioning.

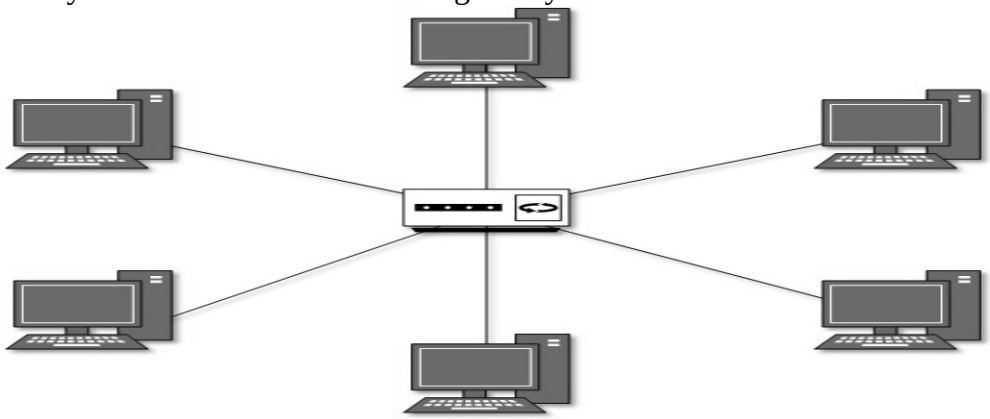


Both ends of the shared channel have line terminator. The data is sent in only one direction and as soon as it reaches the extreme end, the terminator removes the data from the line.

Star Topology

All hosts in Star topology are connected to a central device, known as hub device, using a point-to-point connection. That is, there exists a point to point connection between hosts and hub. The hub device can be any of the following:

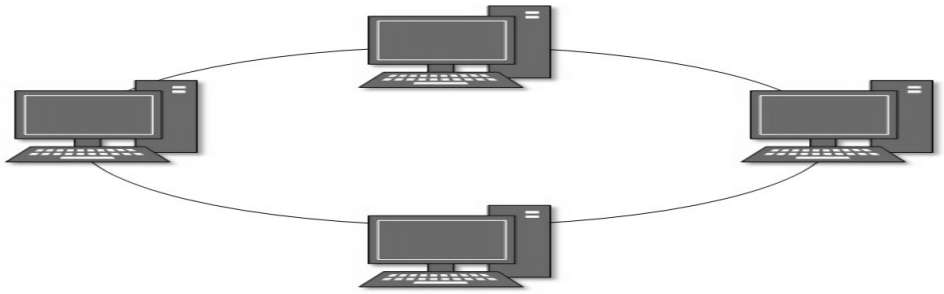
- Layer-1 device such as hub or repeater
- Layer-2 device such as switch or bridge
- Layer-3 device such as router or gateway



As in Bus topology, hub acts as single point of failure. If hub fails, connectivity of all hosts to all other hosts fails. Every communication between hosts takes place through only the hub. Star topology is not expensive as to connect one more host, only one cable is required and configuration is simple.

Ring Topology

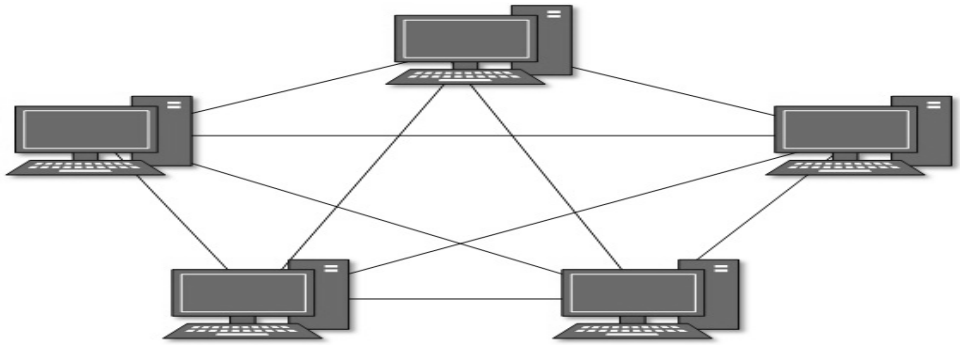
In ring topology, each host machine connects to exactly two other machines, creating a circular network structure. When one host tries to communicate or send message to a host which is not adjacent to it, the data travels through all intermediate hosts. To connect one more host in the existing structure, the administrator may need only one more extra cable.



Failure of any host results in failure of the whole ring. Thus, every connection in the ring is a point of failure. There are methods which employ one more backup ring.

Mesh Topology

In this type of topology, a host is connected to one or multiple hosts. This topology has hosts in point-to-point connection with every other host or may also have hosts which are in point-to-point connection with few hosts only.



Hosts in Mesh topology also work as relay for other hosts which do not have direct point-to-point links. Mesh technology comes into two types:

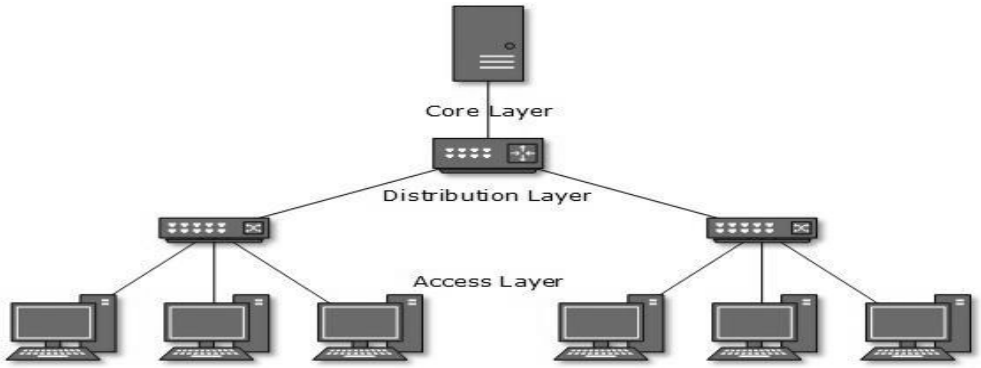
 Full Mesh: All hosts have a point-to-point connection to every other host in the network. Thus for every new host n(n-1)/2 connections are required. It provides the most reliable network structure among all network topologies.

 Partially Mesh: Not all hosts have point-to-point connection to every other host. Hosts connect to each other in some arbitrarily fashion. This topology exists where we need to provide reliability to some hosts out of all.

Tree Topology

Also known as Hierarchical Topology, this is the most common form of network topology in use presently. This topology imitates as extended Star topology and inherits properties of Bus topology.

This topology divides the network into multiple levels/layers of network. Mainly in LANs, a network is bifurcated into three types of network devices. The lowermost is access-layer where computers are attached. The middle layer is known as distribution layer, which works as mediator between upper layer and lower layer. The highest layer is known as core layer, and is central point of the network, i.e. root of the tree from which all nodes fork.



All neighboring hosts have point-to-point connection between them. Similar to the Bus topology, if the root goes down, then the entire network suffers even though it is not the single point of failure. Every connection serves as point of failure, failing of which divides the network into unreachable segment.

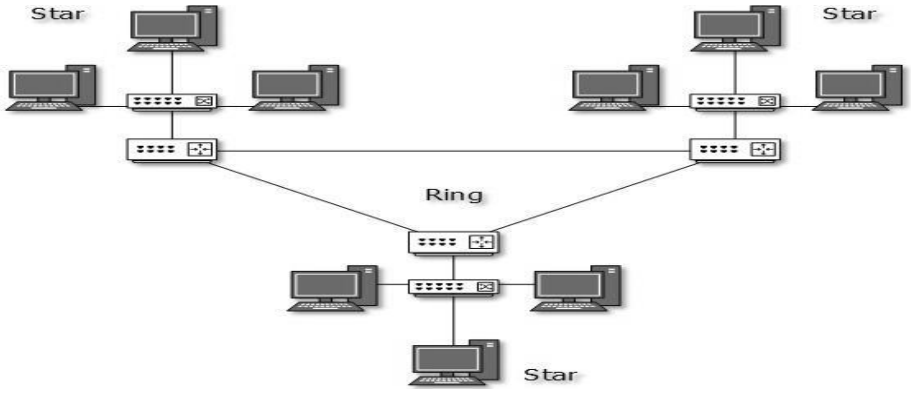
Daisy Chain

This topology connects all the hosts in a linear fashion. Similar to Ring topology, all hosts are connected to two hosts only, except the end hosts. Means, if the end hosts in daisy chain are connected then it represents Ring topology.

Each link in daisy chain topology represents single point of failure. Every link failure splits the network into two segments. Every intermediate host works as relay for its immediate hosts.

Hybrid Topology

A network structure whose design contains more than one topology is said to be hybrid topology. Hybrid topology inherits merits and demerits of all the incorporating topologies.



The above picture represents an arbitrarily hybrid topology. The combining topologies may contain attributes of Star, Ring, Bus, and Daisy-chain topologies. Most WANs are connected by means of Dual-Ring topology and networks connected to them are mostly Star topology networks. Internet is the best example of largest Hybrid topology.

Computer Network Components

- **Servers** - computers that provide shared resources for network users.
- **Clients** - computers that access shared resources provided by servers.
- **Media** - the wires that make the physical connections.
- **Shared data** - files provided to clients by servers across the network.
- **Shared peripherals** - additional hardware resources provided by servers.

Computer Network Models

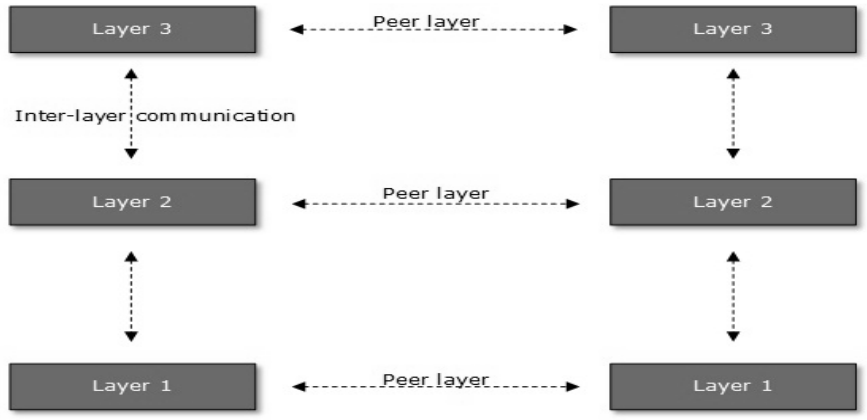
Network engineering is a complicated task, which involves software, firmware, chip level engineering, hardware, and electric pulses. To ease network engineering, the whole networking concept is divided into multiple layers. Each layer is involved in some particular task and is independent of all other layers. But as a whole, almost all networking tasks depend on all of these layers. Layers share data between them and they depend on each other only to take input and send output.

Layered Tasks

In layered architecture of Network Model, one whole network process is divided into small tasks. Each small task is then assigned to a particular layer which works dedicatedly to process the task only. Every layer does only specific work.

In layered communication system, one layer of a host deals with the task done by or to be done by its peer layer at the same level on the remote host. The task is either initiated by layer at the lowest level or at the top most level. If the task is initiated by the topmost layer, it is passed on to

the layer below it for further processing. The lower layer does the same thing, it processes the task and passes on to lower layer. If the task is initiated by lowermost layer, then the reverse path is taken.



Every layer clubs together all procedures, protocols, and methods which it requires to execute its piece of task. All layers identify their counterparts by means of encapsulation header and tail.

OSI Model

Open System Interconnect is an open standard for all communication systems. OSI model is established by International Standard Organization (ISO). This model has seven layers:

- Application Layer
- Presentation Layer
- Session Layer
- Transport Layer
- Network Layer
- Data Link Layer
- Physical Layer

Internet Model

Internet uses TCP/IP protocol suite, also known as Internet suite. This defines Internet Model which contains four layered architecture. OSI Model is general communication model but Internet Model is what the internet uses for all its communication. The internet is independent of its underlying network architecture so is its Model. This model has the following layers:

- Application Layer
- Transport Layer
- Internet Layer
- Link Layer

CHAPTER THREE
DATA COMMUNICATION AND TRANSMISSION MEDIAS

The transmission media is nothing but the physical media over which communication takes place in computer networks.

Magnetic/Guided Media

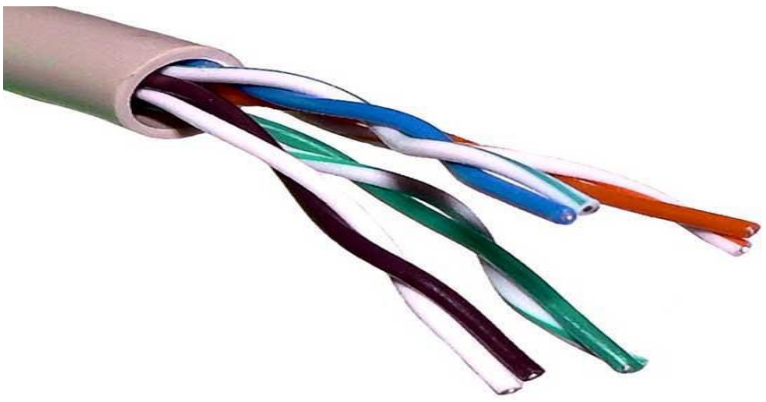
One of the most convenient way to transfer data from one computer to another, even before the birth of networking, was to save it on some storage media and transfer physical from one station to another. Though it may seem old-fashion way in today's world of high speed internet, but when the size of data is huge, the magnetic media comes into play.

For example, a bank has to handle and transfer huge data of its customer, which stores a backup of it at some geographically far-away place for security reasons and to keep it from uncertain calamities. If the bank needs to store its huge backup data, then its transfer through internet is not feasible. The WAN links may not support such high speed. Even if they do; the cost is too high to afford.

In these cases, data backup is stored onto magnetic tapes or magnetic discs, and then shifted physically at remote places.

Twisted Pair Cable

A twisted pair cable is made of two plastic insulated copper wires twisted together to form a single media. Out of these two wires, only one carries actual signal and another is used for ground reference. The twists between wires are helpful in reducing noise (electro-magnetic interference) and crosstalk.



There are two types of twisted pair cables:

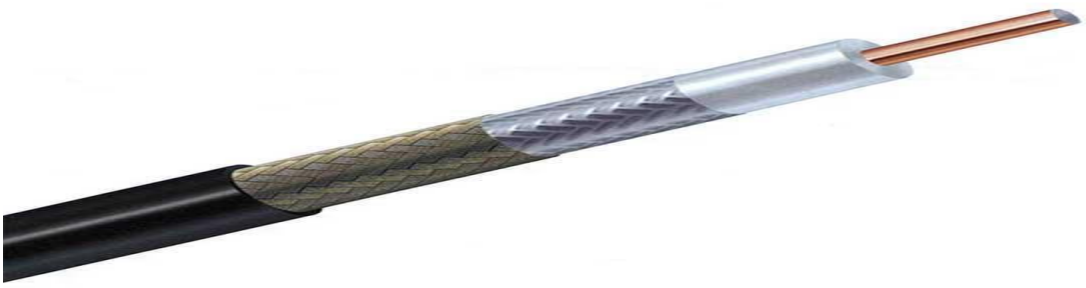
- Shielded Twisted Pair (STP) Cable
- Unshielded Twisted Pair (UTP) Cable

STP cables comes with twisted wire pair covered in metal foil. This makes it more indifferent to noise and crosstalk.

UTP has seven categories, each suitable for specific use. In computer networks, Cat-5, Cat-5e, and Cat-6 cables are mostly used. UTP cables are connected by RJ45 connectors.

Coaxial Cable

Coaxial cable has two wires of copper. The core wire lies in the center and it is made of solid conductor. The core is enclosed in an insulating sheath. The second wire is wrapped around over the sheath and that too in turn encased by insulator sheath. This all is covered by plastic cover.



Because of its structure, the coax cable is capable of carrying high frequency signals than that of twisted pair cable. The wrapped structure provides it a good shield against noise and cross talk. Coaxial cables provide high bandwidth rates of up to 450 mbps.

There are three categories of coax cables namely, RG-59 (Cable TV), RG-58 (Thin Ethernet), and RG-11 (Thick Ethernet). RG stands for Radio Government.

Cables are connected using BNC connector and BNC-T. BNC terminator is used to terminate the wire at the far ends.

Power Lines

Power Line communication (PLC) is Layer-1 (Physical Layer) technology which uses power cables to transmit data signals. In PLC, modulated data is sent over the cables. The receiver on the other end de-modulates and interprets the data.

Because power lines are widely deployed, PLC can make all powered devices controlled and monitored. PLC works in half-duplex.

There are two types of PLCs:

- Narrow band PLC
- Broad band PLC

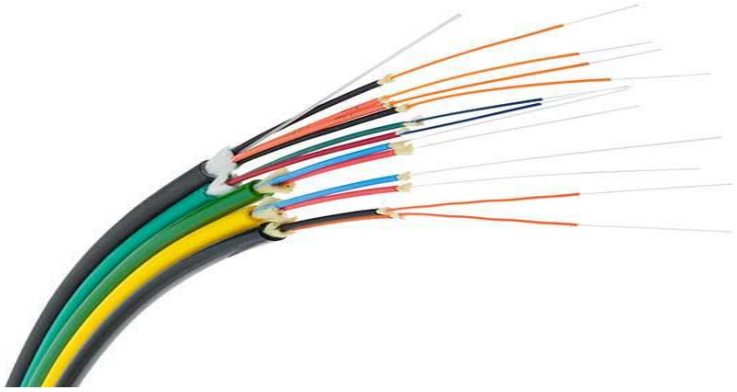
Narrow band PLC provides lower data rates up to 100s of kbps, as they work at lower frequencies (3-5000 kHz). They can be spread over several kilometers.

Broadband PLC provides higher data rates up to 100s of Mbps and works at higher frequencies (1.8 – 250 MHz). They cannot be as much extended as Narrowband PLC.

Fiber Optics

Fiber Optic works on the properties of light. When light ray hits at critical angle, it tends to refracts at 90 degree. This property has been used in fiber optic. The core of fiber optic cable is

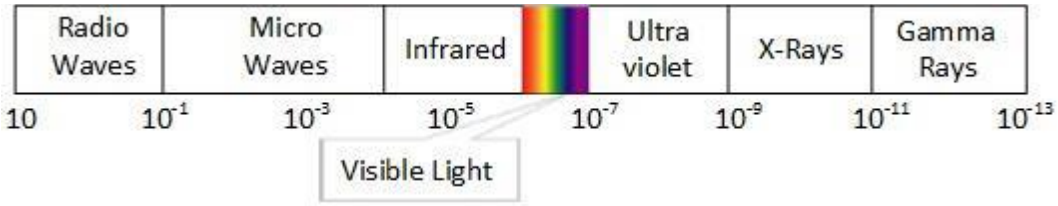
made of high quality glass or plastic. From one end of it light is emitted, it travels through it and at the other end light detector detects light stream and converts it to electric data.
Fiber Optic provides the highest mode of speed. It comes in two modes, one is single mode fiber and second is multimode fiber. Single mode fiber can carry a single ray of light whereas multimode is capable of carrying multiple beams of light.



Fiber Optic also comes in unidirectional and bidirectional capabilities. To connect and access fiber optic special type of connectors are used. These can be Subscriber Channel (SC), Straight Tip (ST), or MT-RJ.

Wireless Transmission

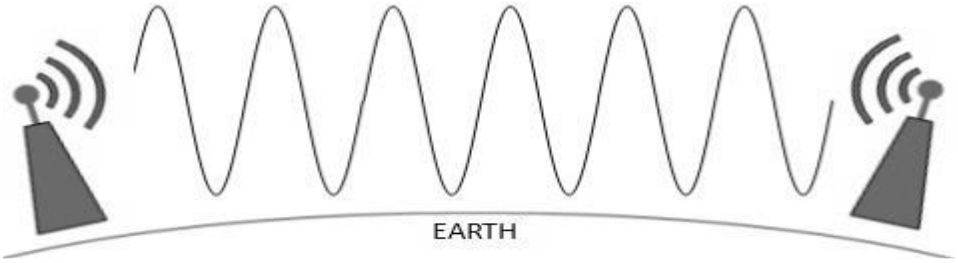
Wireless transmission is a form of unguided media. Wireless communication involves no physical link established between two or more devices, communicating wirelessly. Wireless signals are spread over in the air and are received and interpreted by appropriate antennas.
When an antenna is attached to electrical circuit of a computer or wireless device, it converts the digital data into wireless signals and spread all over within its frequency range. The receptor on the other end receives these signals and converts them back to digital data.
A little part of electromagnetic spectrum can be used for wireless transmission.



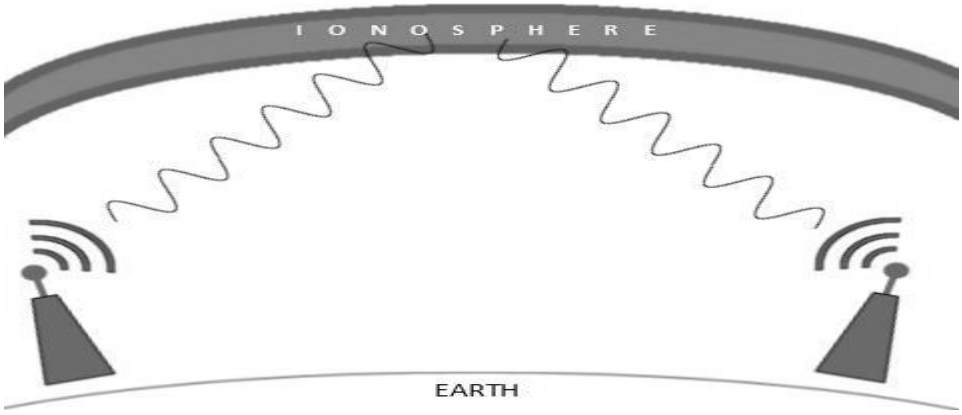
Radio Transmission

Radio frequency is easier to generate and because of its large wavelength it can penetrate through walls and structures alike. Radio waves can have wavelength from 1mm – 100,000km and have frequency ranging from 3Hz (Extremely Low Frequency) to 300 GHz (Extremely High Frequency). Radio frequencies are sub-divided into six bands.

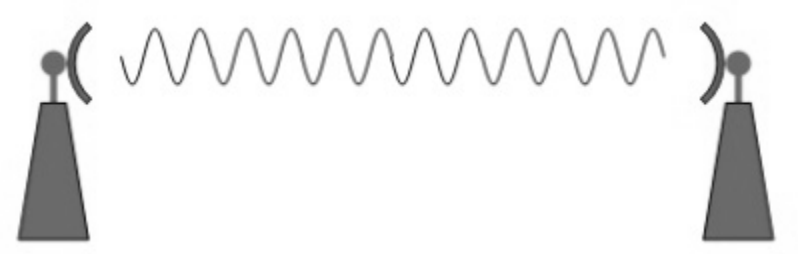
Radio waves at lower frequencies can travel through walls whereas higher RF can travel in straight line and bounce back. The power of low frequency waves decreases sharply as they cover long distance. High frequency radio waves have more power.
Lower frequencies such as VLF, LF, MF bands can travel on the ground up to 1000 kilometers, over the earth's surface.



Radio waves of high frequencies are prone to be absorbed by rain and other obstacles. They use Ionosphere of earth atmosphere. High frequency radio waves such as HF and VHF bands are spread upwards. When they reach Ionosphere, they are refracted back to the earth.



Microwave Transmission
Electromagnetic waves above 100MHz tend to travel in a straight line and signals over them can be sent by beaming those waves towards one particular station. Because Microwaves travels in straight lines, both sender and receiver must be aligned to be strictly in line-of-sight.
Microwaves can have wavelength ranging from 1mm – 1meter and frequency ranging from 300MHz to 300GHz.



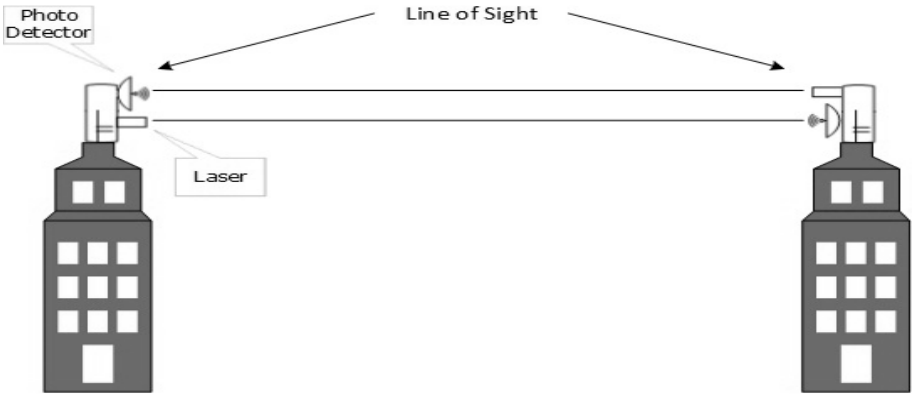
Microwave antennas concentrate the waves making a beam of it. As shown in picture above, multiple antennas can be aligned to reach farther. Microwaves have higher frequencies and do not penetrate wall like obstacles.
Microwave transmission depends highly upon the weather conditions and the frequency it is using.

Infrared Transmission

Infrared wave lies in between visible light spectrum and microwaves. It has wavelength of 700nm to 1mm and frequency ranges from 300GHz to 430THz.
Infrared wave is used for very short range communication purposes such as television and its remote. Infrared travels in a straight line hence it is directional by nature. Because of high frequency range, Infrared cannot cross wall-like obstacles.

Light Transmission

Highest most electromagnetic spectrum which can be used for data transmission is light or optical signaling. This is achieved by means of LASER.
Because of frequency light uses, it tends to travel strictly in straight line. Hence the sender and receiver must be in the line-of-sight. Because laser transmission is unidirectional, at both ends of communication the laser and the photo-detector needs to be installed. Laser beam is generally 1mm wide hence it is a work of precision to align two far receptors each pointing to lasers source.



Laser works as Tx (transmitter) and photo-detectors works as Rx (receiver).

Lasers cannot penetrate obstacles such as walls, rain, and thick fog. Additionally, laser beam is distorted by wind, atmosphere temperature, or variation in temperature in the path.
Laser is safe for data transmission as it is very difficult to tap 1mm wide laser without interrupting the communication channel.

CHAPTER FOUR

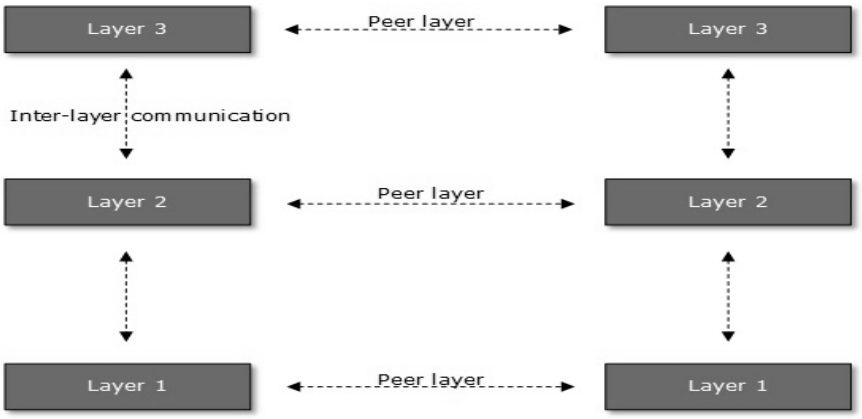
LAYERED MODELS

Network engineering is a complicated task, which involves software, firmware, chip level engineering, hardware, and electric pulses. To ease network engineering, the whole networking concept is divided into multiple layers. Each layer is involved in some particular task and is independent of all other layers. But as a whole, almost all networking tasks depend on all of these layers. Layers share data between them and they depend on each other only to take input and send output.

Layered Architecture

In layered architecture of Network Model, one whole network process is divided into small tasks. Each small task is then assigned to a particular layer which works dedicatedly to process the task only. Every layer does only specific work.

In layered communication system, one layer of a host deals with the task done by or to be done by its peer layer at the same level on the remote host. The task is either initiated by layer at the lowest level or at the top most level. If the task is initiated by the topmost layer, it is passed on to the layer below it for further processing. The lower layer does the same thing, it processes the task and passes on to lower layer. If the task is initiated by lowermost layer, then the reverse path is taken.



Every layer clubs together all procedures, protocols, and methods which it requires to execute its piece of task. All layers identify their counterparts by means of encapsulation header and tail.

THE OSI MODEL

Open System Interconnect is an open standard for all communication systems. OSI model is established by International Standard Organization (ISO). The OSI model is not a protocol; it is a model for understanding and designing a network architecture that is flexible, robust, and

interoperable. The OSI model is a layered framework for the design of network systems that allows communication between all types of computer systems. This model has seven layers:

Application Layer: This layer is responsible for providing interface to the application user. This layer encompasses protocols which directly interact with the user.

Presentation Layer: This layer defines how data in the native format of remote host should be presented in the native format of host.

Session Layer: This layer maintains sessions between remote hosts. For example, once user/password authentication is done, the remote host maintains this session for a while and does not ask for authentication again in that time span.

Transport Layer: This layer is responsible for end-to-end delivery between hosts.

Network Layer: This layer is responsible for address assignment and uniquely addressing hosts in a network.

Data Link Layer: This layer is responsible for reading and writing data from and onto the line. Link errors are detected at this layer.

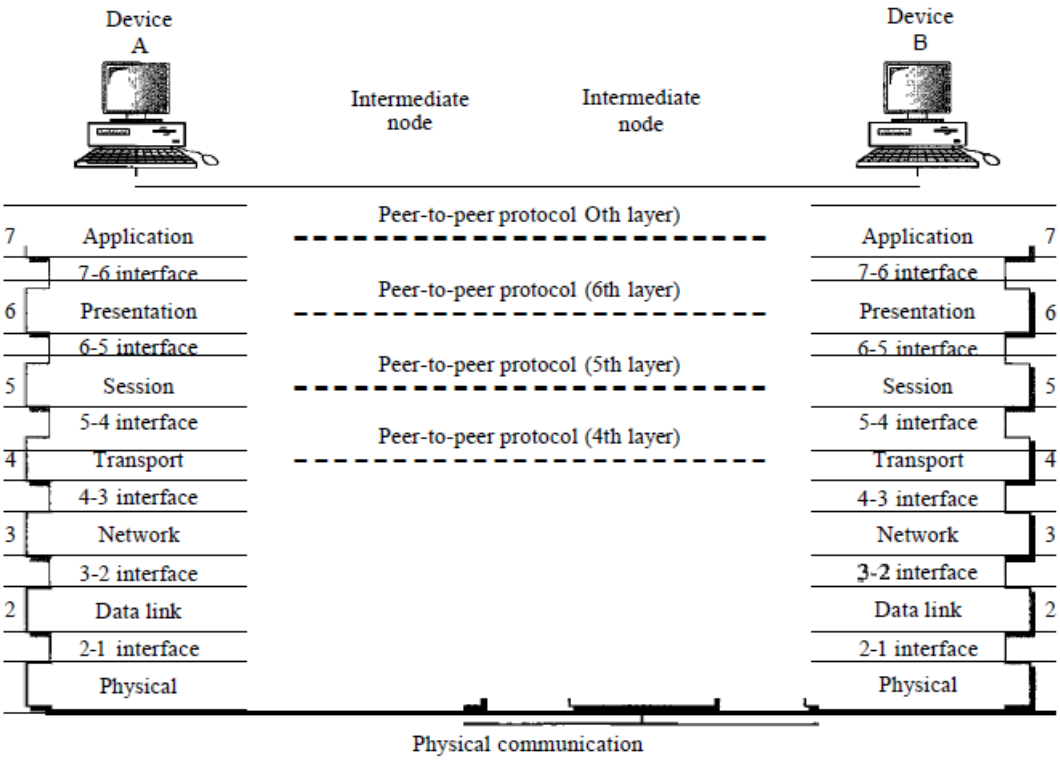
Physical Layer: This layer defines the hardware, cabling, wiring, power output, pulse rate etc.

Peer-to-Peer Processes

The processes on each machine that communicate at a given layer are called peer-to-peer processes. Communication between machines is therefore a peer-to-peer process using the protocols appropriate to a given layer.

At the physical layer, communication is direct: In Figure below, device A sends a stream of bits to device B (through intermediate nodes). At the higher layers, however, communication must move down through the layers on device A, over to device B, and then back up through the layers. Each layer in the sending device adds its own information to the message it receives from the layer just above it and passes the whole package to the layer just below it.

At layer I the entire package is converted to a form that can be transmitted to the receiving device. At the receiving machine, the message is unwrapped layer by layer, with each process receiving and removing the data meant for it. For example, layer 2 removes the data meant for it, then passes the rest to layer 3. Layer 3 then removes the data meant for it and passes the rest to layer 4, and so on.



Encapsulation

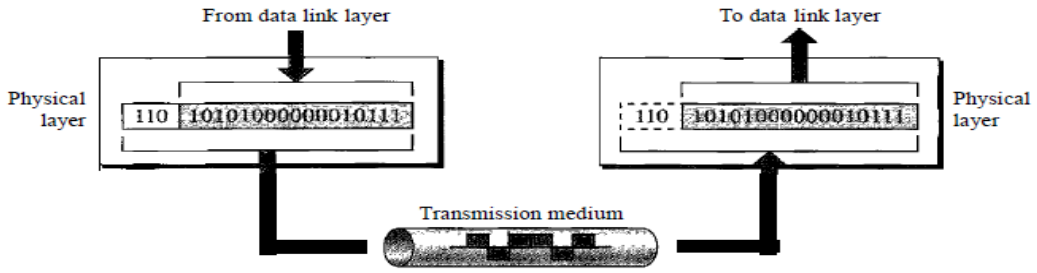
Figure described above (under peer-to-peer process) reveals another aspect of data communications in the OSI model: encapsulation. A packet (header and data) at level 7 is encapsulated in a packet at level 6. The whole packet at level 6 is encapsulated in a packet at level 5, and so on.

In other words, the data portion of a packet at level *N* - 1 carries the whole packet (data and header and maybe trailer) from level *N.* The concept is called *encapsulation;* level *N* - 1 is not aware of which part of the encapsulated packet is data and which part is the header or trailer. For level *N* - 1, the whole packet coming from level *N* is treated as one integral unit.

Physical layer

The physical layer coordinates the functions required to carry a bit stream over a physical medium. It deals with the mechanical and electrical specifications of the interface and transmission medium. It also defines the procedures and functions that physical devices and interfaces have to perform for transmission to Occur.

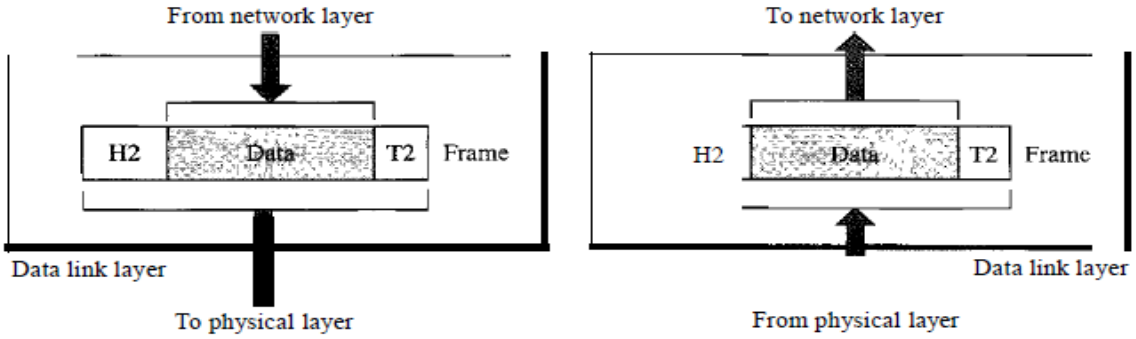
Figure shows the position of the physical layer with respect to the transmission medium and the data link layer.



The physical layer is responsible for movements of individual bits from one hop (node) to the next.

Data-link Layer

The data link layer transforms the physical layer, a raw transmission facility, to a reliable link. It makes the physical layer appear error-free to the upper layer (network layer). Figure below shows the relationship of the data link layer to the network and physical layers.



The data link layer is responsible for moving frames from one hop (node) to the next.

Other responsibilities of the data link layer include the following:

- **Framing**. The data link layer divides the stream of bits received from the network layer into manageable data units called frames.
- **Physical addressing.** If frames are to be distributed to different systems on the network, the data link layer adds a header to the frame to define the sender and/or receiver of the frame. If the frame is intended for a system outside the sender's network, the receiver address is the address of the device that connects the network to the next one.
- **Flow control.** If the rate at which the data are absorbed by the receiver is less than the rate at which data are produced in the sender, the data link layer imposes a flow control mechanism to avoid overwhelming the receiver.
- **Error control.** The data link layer adds reliability to the physical layer by adding mechanisms to detect and retransmit damaged or lost frames. It also uses a mechanism to

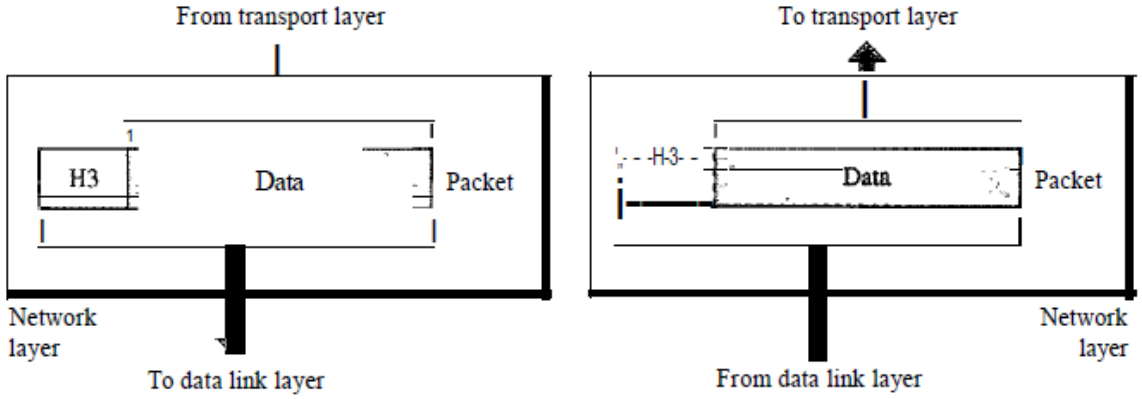
recognize duplicate frames. Error control is normally achieved through a trailer added to the end of the frame.

- **Access control.** When two or more devices are connected to the same link, data link layer protocols are necessary to determine which device has control over the link at any given time.

Network Layer and Routing

The network layer is responsible for the source-to-destination delivery of a packet, possibly across multiple networks (links). Whereas the data link layer oversees the delivery of the packet between two systems on the same network (links), the network layer ensures that each packet gets from its point of origin to its final destination.

If two systems are connected to the same link, there is usually no need for a network layer. However, if the two systems are attached to different networks (links) with connecting devices between the networks (links), there is often a need for the network layer to accomplish source-to-destination delivery. Figure below shows the relationship of the network layer to the data link and transport layers.



The network layer is responsible for the delivery of individual packets from the source host to the destination host.

Other responsibilities of the network layer include the following:

Logical addressing. The physical addressing implemented by the data link layer handles the addressing problem locally. If a packet passes the network boundary, we need another addressing system to help distinguish the source and destination systems. The network layer adds a header to the packet coming from the upper layer that, among other things, includes the logical addresses of the sender and receiver.

Routing. When independent networks or links are connected to create *internetworks* (network of networks) or a large network, the connecting devices (called *routers* or *switches)* route or switch the packets to their final destination. One of the functions of the network layer is to provide this mechanism.

Network Addressing

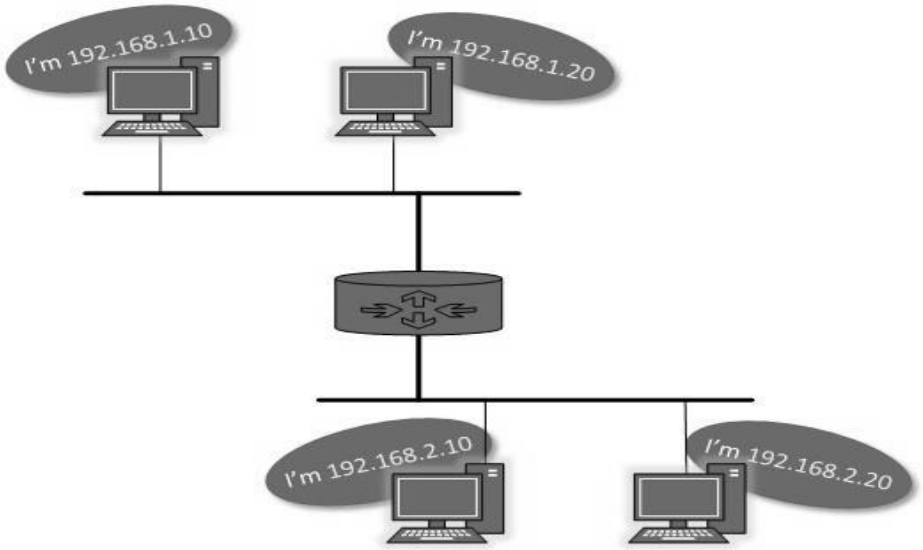
Layer 3 network addressing is one of the major tasks of Network Layer. Network Addresses are always logical i.e. these are software based addresses which can be changed by appropriate configurations.

A network address always points to host / node / server or it can represent a whole network. Network address is always configured on network interface card and is generally mapped by system with the MAC address (hardware address or layer-2 address) of the machine for Layer-2 communication.

There are different kinds of network addresses in existence:

- IP
- IPX
- AppleTalk

We are discussing IP here as it is the only one we use in practice these days.



IP addressing provides mechanism to differentiate between hosts and network. Because IP addresses are assigned in hierarchical manner, a host always resides under a specific network. The host which needs to communicate outside its subnet, needs to know destination network address, where the packet/data is to be sent.

Hosts in different subnet need a mechanism to locate each other. This task can be done by DNS. DNS is a server which provides Layer-3 address of remote host mapped with its domain name or FQDN. When a host acquires the Layer-3 Address (IP Address) of the remote host, it forwards all its packet to its gateway. A gateway is a router equipped with all the information which leads to route packets to the destination host.

Routers take help of routing tables, which has the following information:

- Address of destination network
- Method to reach the network

Routers upon receiving a forwarding request, forwards packet to its next hop (adjacent router) towards the destination.
The next router on the path follows the same thing and eventually the data packet reaches its destination.
Network address can be of one of the following:

- Unicast (destined to one host)
- Multicast (destined to group)
- Broadcast (destined to all)
- Anycast (destined to nearest one)

A router never forwards broadcast traffic by default. Multicast traffic uses special treatment as it is most a video stream or audio with highest priority. Anycast is just similar to unicast, except that the packets are delivered to the nearest destination when multiple destinations are available.

Network Routing

When a device has multiple paths to reach a destination, it always selects one path by preferring it over others. This selection process is termed as Routing. Routing is done by special network devices called routers or it can be done by means of software processes. The software based routers have limited functionality and limited scope.
A router is always configured with some default route. A default route tells the router where to forward a packet if there is no route found for specific destination.

Routes can be statically configured or dynamically learnt. One route can be configured to be preferred over others.

Unicast routing
Most of the traffic on the internet and intranets known as unicast data or unicast traffic is sent with specified destination. Routing unicast data over the internet is called unicast routing. It is the simplest form of routing because the destination is already known. Hence the router just has to look up the routing table and forward the packet to next hop.

Broadcast routing
By default, the broadcast packets are not routed and forwarded by the routers on any network. Routers create broadcast domains. But it can be configured to forward broadcasts in some special cases. A broadcast message is destined to all network devices.
Broadcast routing can be done in two ways (algorithm):

- A router creates a data packet and then sends it to each host one by one. In this case, the router creates multiple copies of single data packet with different destination addresses. All packets are sent as unicast but because they are sent to all, it simulates as if router is broadcasting.

- This method consumes lots of bandwidth and router must destination address of each node.
- Secondly, when router receives a packet that is to be broadcasted, it simply floods those packets out of all interfaces. All routers are configured in the same way.

Multicast Routing

Multicast routing is special case of broadcast routing with significance difference and challenges. In broadcast routing, packets are sent to all nodes even if they do not want it. But in Multicast routing, the data is sent to only nodes which wants to receive the packets.

The router must know that there are nodes, which wish to receive multicast packets (or stream) then only it should forward. Multicast routing works spanning tree protocol to avoid looping.
Multicast routing also uses reverse path Forwarding technique, to detect and discard duplicates and loops.

Anycast Routing

Anycast packet forwarding is a mechanism where multiple hosts can have same logical address. When a packet destined to this logical address is received, it is sent to the host which is nearest in routing topology.

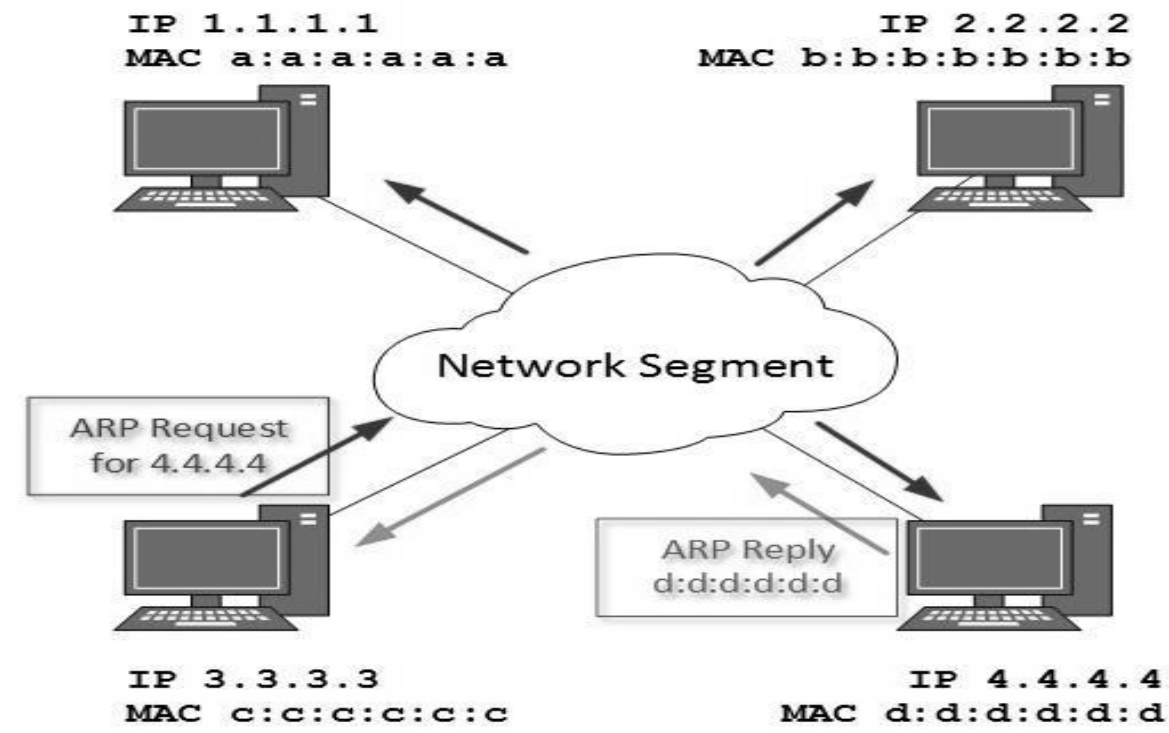
Anycast routing is done with help of DNS server. Whenever an Anycast packet is received it is enquired with DNS to where to send it. DNS provides the IP address which is the nearest IP configured on it.

Network Layer Protocols

Every computer in a network has an IP address by which it can be uniquely identified and addressed. An IP address is Layer-3 (Network Layer) logical address. This address may change every time a computer restarts. A computer can have one IP at one instance of time and another IP at some different time.

Address Resolution Protocol (ARP)

While communicating, a host needs Layer-2 (MAC) address of the destination machine which belongs to the same broadcast domain or network. A MAC address is physically burnt into the Network Interface Card (NIC) of a machine and it never changes.
On the other hand, IP address on the public domain is rarely changed. If the NIC is changed in case of some fault, the MAC address also changes. This way, for Layer-2 communication to take place, a mapping between the two is required.



To know the MAC address of remote host on a broadcast domain, a computer wishing to initiate communication sends out an ARP broadcast message asking, "Who has this IP address?" Because it is a broadcast, all hosts on the network segment (broadcast domain) receive this packet and process it. ARP packet contains the IP address of destination host, the sending host wishes to talk to. When a host receives an ARP packet destined to it, it replies back with its own MAC address.

Once the host gets destination MAC address, it can communicate with remote host using Layer-2 link protocol. This MAC to IP mapping is saved into ARP cache of both sending and receiving hosts. Next time, if they require to communicate, they can directly refer to their respective ARP cache.

Reverse ARP is a mechanism where host knows the MAC address of remote host but requires to know IP address to communicate.

Internet Control Message Protocol (ICMP)

ICMP is network diagnostic and error reporting protocol. ICMP belongs to IP protocol suite and uses IP as carrier protocol. After constructing ICMP packet, it is encapsulated in IP packet. Because IP itself is a best-effort non-reliable protocol, so is ICMP.

Any feedback about network is sent back to the originating host. If some error in the network occurs, it is reported by means of ICMP. ICMP contains dozens of diagnostic and error reporting messages.

ICMP-echo and ICMP-echo-reply are the most commonly used ICMP messages to check the reachability of end-to-end hosts. When a host receives an ICMP-echo request, it is bound to send

back an ICMP-echo-reply. If there is any problem in the transit network, the ICMP will report that problem.

Internet Protocol Version 4 (IPv4)

IPv4 is 32-bit addressing scheme used as TCP/IP host addressing mechanism. IP addressing enables every host on the TCP/IP network to be uniquely identifiable.

IPv4 provides hierarchical addressing scheme which enables it to divide the network into sub-networks, each with well-defined number of hosts. IP addresses are divided into many categories:

- **Class A:** It uses first octet for network addresses and last three octets for host addressing.
- **Class B:** It uses first two octets for network addresses and last two for host addressing.
- **Class C:** It uses first three octets for network addresses and last one for host addressing.
- **Class D:** It provides flat IP addressing scheme in contrast to hierarchical structure for above three.
- **Class E:** It is used as experimental.

IPv4 also has well-defined address spaces to be used as private addresses (not routable on internet), and public addresses (provided by ISPs and are routable on internet).

Though IP is not reliable one; it provides 'Best-Effort-Delivery' mechanism.

Internet Protocol Version 6 (IPv6)

Exhaustion of IPv4 addresses gave birth to a next generation Internet Protocol version 6. IPv6 addresses its nodes with 128-bit wide address providing plenty of address space for future to be used on entire planet or beyond.

IPv6 has introduced Anycast addressing but has removed the concept of broadcasting. IPv6 enables devices to self-acquire an IPv6 address and communicate within that subnet. This auto-configuration removes the dependability of Dynamic Host Configuration Protocol (DHCP) servers. This way, even if the DHCP server on that subnet is down, the hosts can communicate with each other.

IPv6 provides new feature of IPv6 mobility. Mobile IPv6-equipped machines can roam around without the need of changing their IP addresses.

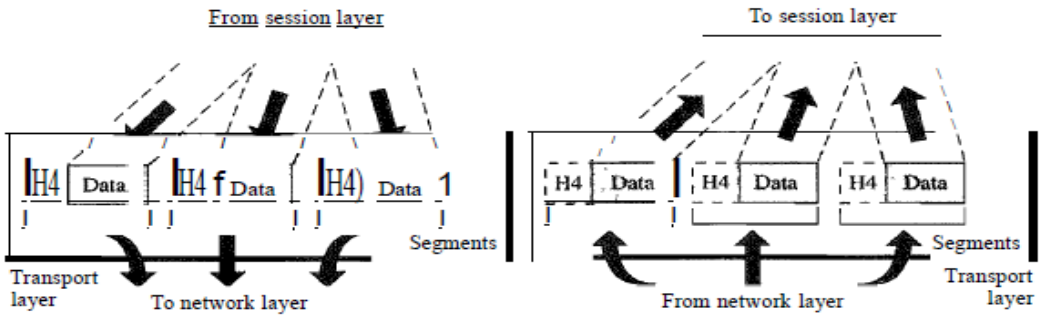
IPv6 is still in transition phase and is expected to replace IPv4 completely in coming years. At present, there are few networks which are running on IPv6. There are some transition mechanisms available for IPv6-enabled networks to speak and roam around different networks easily on IPv4. These are:

- Dual stack implementation
- Tunneling
- NAT-PT

Transport Layer

The transport layer is responsible for process-to-process delivery of the entire message. A process is an application program running on a host. Whereas the network layer oversees source-to-destination delivery of individual packets, it does not recognize any relationship between

those packets. It treats each one independently, as though each piece belonged to a separate message, whether or not it does. The transport layer, on the other hand, ensures that the whole message arrives intact and in order, overseeing both error control and flow control at the source-to-destination level.



The transport layer is responsible for the delivery of a message from one process to another.

Other responsibilities of the transport layer include the following:

Service-point addressing. Computers often run several programs at the same time. For this reason, source-to-destination delivery means delivery not only from one computer to the next but also from a specific process (running program) on one computer to a specific process (running program) on the other. The transport layer header must therefore include a type of address called a *service-point address* (or port address). The network layer gets each packet to the correct computer; the transport layer gets the entire message to the correct process on that computer.

Segmentation and reassembly. A message is divided into transmittable segments, with each segment containing a sequence number. These numbers enable the transport layer to reassemble the message correctly upon arriving at the destination and to identify and replace packets that were lost in transmission.

Connection control. The transport layer can be either connectionless or connection-oriented.

A connectionless transport layer treats each segment as an independent packet and delivers it to the transport layer at the destination machine. A connection-oriented transport layer makes a connection with the transport layer at the destination machine first before delivering the packets. After all the data are transferred, the connection is terminated.

Flow control. Like the data link layer, the transport layer is responsible for flow control. However, flow control at this layer is performed end to end rather than across a single link.

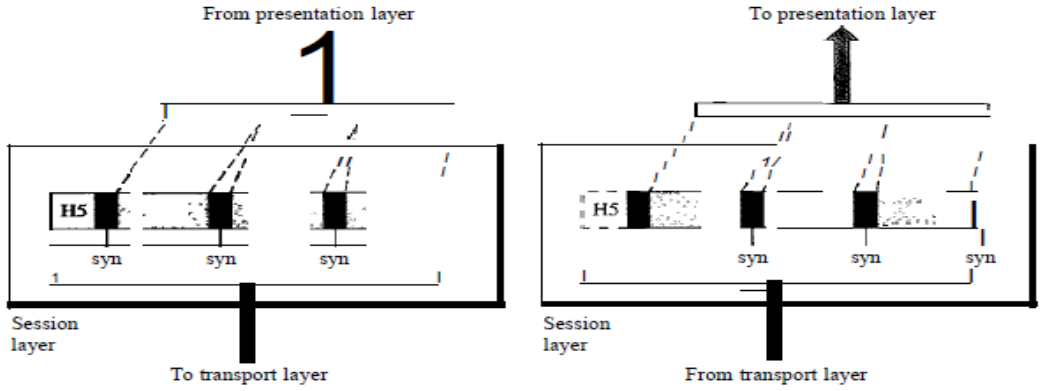
Error control. Like the data link layer, the transport layer is responsible for error control. However, error control at this layer is performed process-to process rather than across a single link. The sending transport layer makes sure that the entire message arrives at the receiving transport layer without error (damage, loss, or duplication). Error correction is usually achieved through retransmission.

Session Layer

The services provided by the first three layers (physical, data link, and network) are not sufficient for some processes. The session layer is the network *dialog controller*. It establishes, maintains, and synchronizes the interaction among communicating systems.

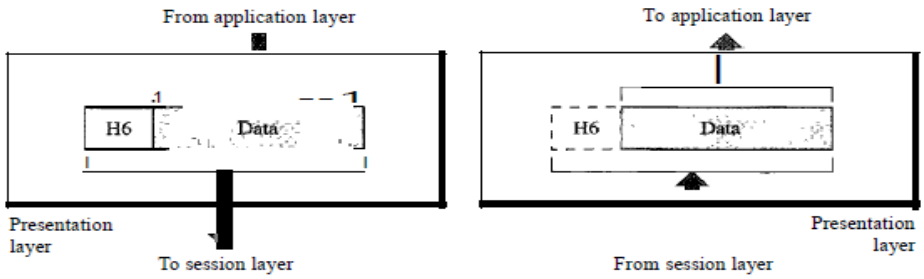
The session layer is responsible for dialog control and synchronization.

Specific responsibilities of the session layer include the following:
Dialog control. The session layer allows two systems to enter into a dialog. It allows the communication between two processes to take place in either half-duplex (one way at a time) or full-duplex (two ways at a time) mode.
Synchronization. The session layer allows a process to add checkpoints, or synchronization points, to a stream of data. For example, if a system is sending a file of 2000 pages, it is advisable to insert checkpoints after every 100 pages to ensure that each 100-page unit is received and acknowledged independently. In this case, if a crash happens during the transmission of page 523, the only pages that need to be resent after system recovery are pages 501 to 523. Pages previous to 501 need not be resent. Figure below illustrates the relationship of the session layer to the transport and presentation layers.



Presentation Layer

The presentation layer is concerned with the syntax and semantics of the information exchanged between two systems. Figure below shows the relationship between the presentation layer and the application and session layers.



The presentation layer is responsible for translation, compression, and encryption.

Specific responsibilities of the presentation layer include the following:

Translation. The processes (running programs) in two systems are usually exchanging information in the form of character strings, numbers, and so on. The information must be changed to bit streams before being transmitted. Because different computers use different encoding systems, the presentation layer is responsible for interoperability between these different encoding methods. The presentation layer at the sender changes the information from its sender-dependent format into a common format. The presentation layer at the receiving machine changes the common format into its receiver-dependent format.

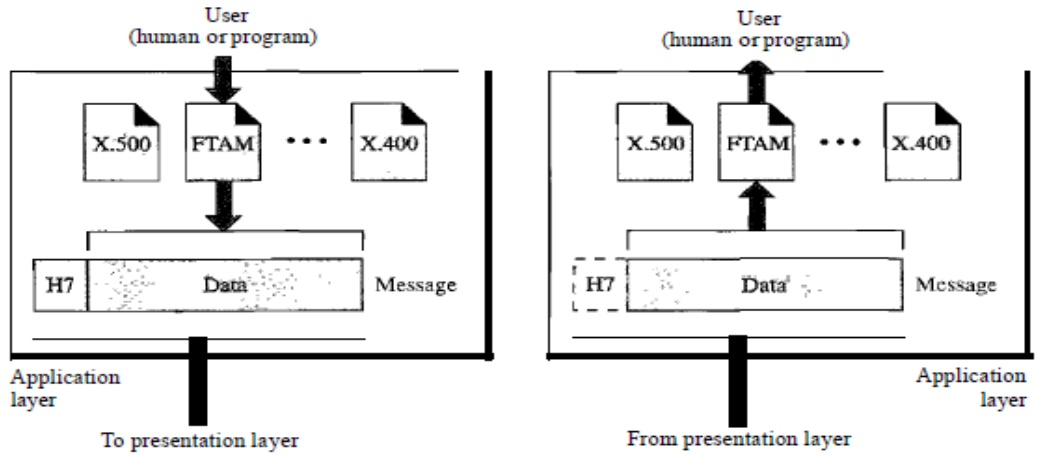
Encryption. To carry sensitive information, a system must be able to ensure privacy. Encryption means that the sender transforms the original information to another form and sends the resulting message out over the network. Decryption reverses the original process to transform the message back to its original form.

Compression. Data compression reduces the number of bits contained in the information. Data compression becomes particularly important in the transmission of multimedia such as text, audio, and video.

Application Layer

The application layer enables the user, whether human or software, to access the network. It provides user interfaces and support for services such as electronic mail, remote file access and transfer, shared database management, and other types of distributed information services.

Figure 2.14 shows the relationship of the application layer to the user and the presentation layer. Of the many application services available, the figure shows only three: *XAOO* (message-handling services), X.500 (directory services), and file transfer, access, and management (FTAM). The user in this example employs *XAOO* to send an e-mail message.



The application layer is responsible for providing services to the user.

Specific services provided by the application layer include the following:

Network virtual terminal. A network virtual terminal is a software version of a physical terminal, and it allows a user to log on to a remote host. To do so, the application creates a software emulation of a terminal at the remote host. The user's computer talks to the software terminal which, in turn, talks to the host, and vice versa. The remote host believes it is communicating with one of its own terminals and allows the user to log on.

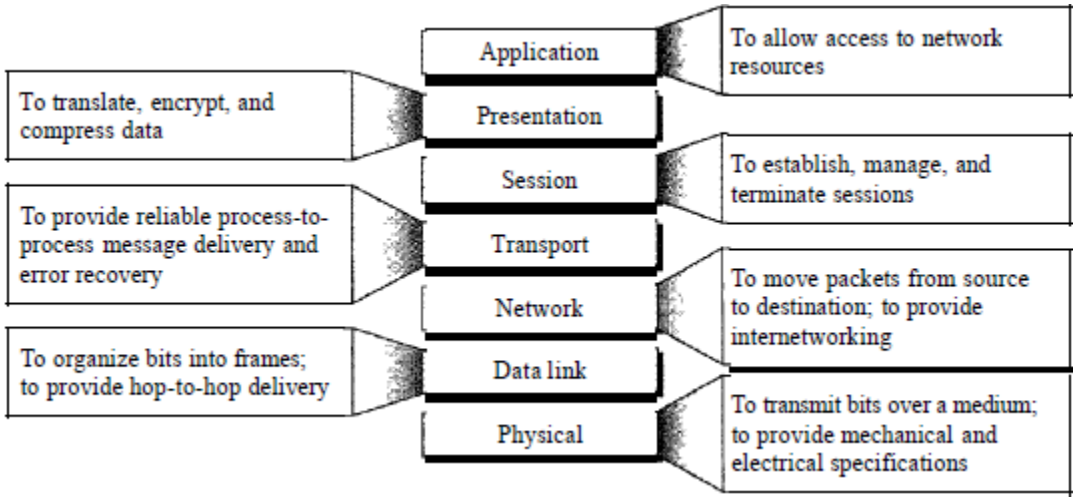
File transfer, access, and management. This application allows a user to access files in a remote host (to make changes or read data), to retrieve files from a remote

Computer for use in the local computer, and to manage or control files in a remote computer locally.

Mail services. This application provides the basis for e-mail forwarding and storage.

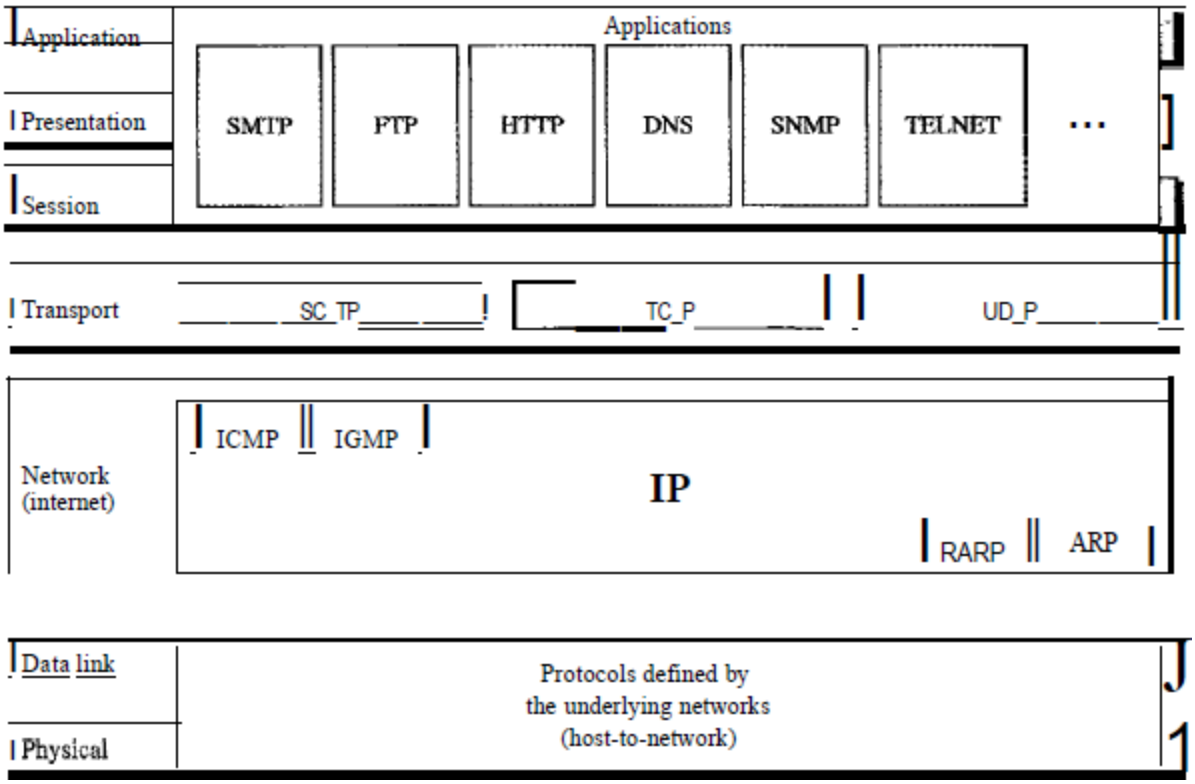
Directory services. This application provides distributed database sources and access for global information about various objects and services.

Summary of Layers



TCP/IP PROTOCOL SUITE

The TCP/IP protocol suite was developed prior to the OSI model. Therefore, the layers in the TCP/IP protocol suite do not exactly match those in the OSI model. The original TCP/IP protocol suite was defined as having four layers: host-to-network, internet, transport, and application. However, when TCP/IP is compared to OSI, we can say that the host-to-network layer is equivalent to the combination of the physical and data link layers. The internet layer is equivalent to the network layer, and the application layer is roughly doing the job of the session, presentation, and application layers with the transport layer in TCPIIP taking care of part of the duties of the session layer. So in this book, we assume that the TCPIIP protocol suite is made of five layers: physical, data link, network, transport, and application. The first four layers provide physical standards, network interfaces, internetworking, and transport functions that correspond to the first four layers of the OSI model. The three topmost layers in the OSI model, however, are represented in TCP/IP by a single layer called the *application layer*.



TCP/IP is a hierarchical protocol made up of interactive modules, each of which provides a specific functionality; however, the modules are not necessarily interdependent.

Whereas the OSI model specifies which functions belong to each of its layers, the layers of the *TCP/IP* protocol suite contain relatively independent protocols that can be mixed and matched depending on the needs of the system. The term *hierarchical* means that each upper-level protocol is supported by one or more lower-level protocols.

At the transport layer, *TCP/IP* defines three protocols: Transmission Control Protocol (TCP), User Datagram Protocol (UDP), and Stream Control Transmission Protocol (SCTP). At the network layer, the main protocol defined by TCP/IP is the Internetworking Protocol (IP); there are also some other protocols that support data movement in this layer.

Physical and Data Link Layers

At the physical and data link layers, *TCP/IP* does not define any specific protocol. It supports all the standard and proprietary protocols. A network in a *TCP/IP* internetwork can be a local-area network or a wide-area network.

Network Layer

At the network layer (or, more accurately, the internetwork layer), *TCP/IP* supports the Internetworking Protocol. IP, in turn, uses four supporting protocols: ARP, RARP, ICMP, and IGMP.

Internetworking Protocol (IP)

The Internetworking Protocol (IP) is the transmission mechanism used by the TCP/IP protocols. It is an unreliable and connectionless protocol-a best-effort delivery service.

The term *best effort* means that IP provides no error checking or tracking. IP assumes the unreliability of the underlying layers and does its best to get a transmission through to its destination, but with no guarantees.

IP transports data in packets called *datagrams*, each of which is transported separately.

Datagrams can travel along different routes and can arrive out of sequence or be duplicated. IP does not keep track of the routes and has no facility for reordering datagrams once they arrive at their destination.

The limited functionality of IP should not be considered a weakness, however. IP provides bare-bones transmission functions that free the user to add only those facilities necessary for a given application and thereby allows for maximum efficiency

Address Resolution Protocol

The Address Resolution Protocol (ARP) is used to associate a logical address with a physical address. On a typical physical network, such as a LAN, each device on a link is identified by a physical or station address, usually imprinted on the network interface card (NIC). ARP is used to find the physical address of the node when its Internet address is known.

Reverse Address Resolution Protocol

The Reverse Address Resolution Protocol (RARP) allows a host to discover its Internet address when it knows only its physical address. It is used when a computer is connected to a network for the first time or when a diskless computer is booted.

Internet Control Message Protocol

The Internet Control Message Protocol (ICMP) is a mechanism used by hosts and gateways to send notification of datagram problems back to the sender. ICMP sends query and error reporting messages.

Internet Group Message Protocol

The Internet Group Message Protocol (IGMP) is used to facilitate the simultaneous transmission of a message to a group of recipients.

Transport Layer

Traditionally the transport layer was represented in *TCP/IP* by two protocols: TCP and UDP. IP is a host-to-host protocol, meaning that it can deliver a packet from one physical device to another. UDP and TCP are transport level protocols responsible for delivery of a message from a process (running program) to another process. A new transport layer protocol, SCTP, has been devised to meet the needs of some newer applications.

User Datagram Protocol

The User Datagram Protocol (UDP) is the simpler of the two standard TCP/IP transport protocols. It is a process-to-process protocol that adds only port addresses, checksum error control, and length information to the data from the upper layer.

Transmission Control Protocol

The Transmission Control Protocol (TCP) provides full transport-layer services to applications. TCP is a reliable stream transport protocol. The term *stream*, in this context, means connection-oriented: A connection must be established between both ends of a transmission before either can transmit data.

At the sending end of each transmission, TCP divides a stream of data into smaller units called *segments*. Each segment includes a sequence number for reordering after receipt, together with an acknowledgment number for the segments received. Segments are carried across the internet inside of IP datagrams. At the receiving end, TCP collects each datagram as it comes in and reorders the transmission based on sequence numbers. TCP is discussed in Chapter 23.

Stream Control Transmission Protocol

The Stream Control Transmission Protocol (SCTP) provides support for newer applications such as voice over the Internet. It is a transport layer protocol that combines the best features of UDP and TCP.

Application Layer

The *application layer* in TCP/IP is equivalent to the combined session, presentation, and application layers in the OSI model. Many protocols are defined at this layer.

Simple Mail Transfer Protocol (SMTP)

Governs the transmission of mail messages and attachments. SMTP is used in the case of outgoing messages. More powerful protocols such as POP3 and IMAP4 are needed and available to manage incoming messages.

- POP3(Post Office Protocol version 3) is the older protocol
- IMAP4(Internet Mail Access Protocol version 4) is the more advanced protocol

Telnet

Telnet is a protocol used to log on to remote hosts using the TCP/IP protocol suite. Using Telnet, a TCP connection is established and keystrokes on the user's machine act like keystrokes on the remotely connected machine. Often, Telnet is used to connect two dissimilar systems (such as PCs and UNIX machines).

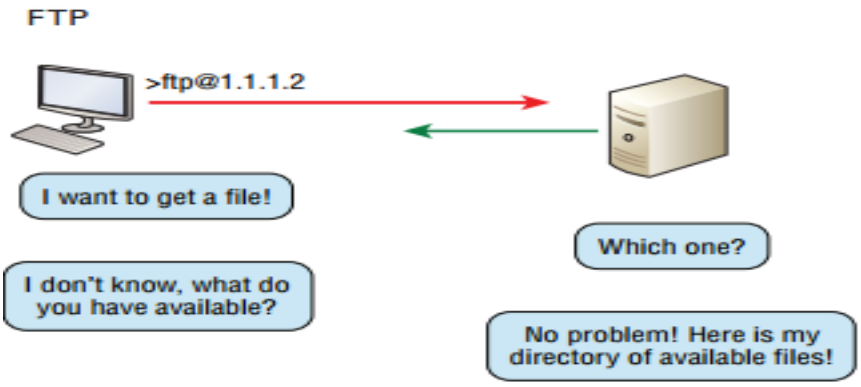
Through Telnet, you can control a remote host over LANs and WANs such as the Internet. For example, network managers can use Telnet to log on to a router from a computer elsewhere on their LAN and modify the router's configuration.

File Transfer Protocol (FTP)

File Transfer Protocol (FTP) lets us transfer files, and it can accomplish this between any two machines using it. But accessing a host through FTP is only the first step. Users must then be

subjected to an authentication login that's usually secured with passwords and usernames implemented by system administrators to restrict access.

FTP's functions are limited to listing and manipulating directories, typing file contents, and copying files between hosts.



Trivial File Transfer Protocol (TFTP)

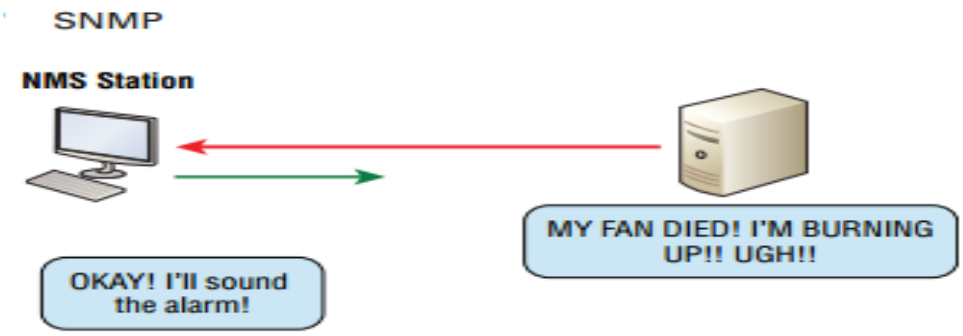
Trivial File Transfer Protocol (TFTP) is stock version of FTP, but it's the protocol of choice if you know exactly what you want and where to find it because it's fast and so easy to use! But TFTP doesn't offer the abundance of functions that FTP does because it has no directory-browsing abilities, meaning that it can only send and receive files. There's no authentication as with FTP, so it's even more insecure, and few sites support it because of the inherent security risks.

A significant difference between FTP and TFTP is that TFTP relies on UDP at the Transport layer, but FTP uses TCP protocol.

Simple Network Management Protocol (SNMP)

Simple Network Management Protocol (SNMP) collects and manipulates valuable network information. It gathers data from a network management station (NMS) at fixed or random intervals, requiring them to disclose certain information, or even asking for certain information from the device.

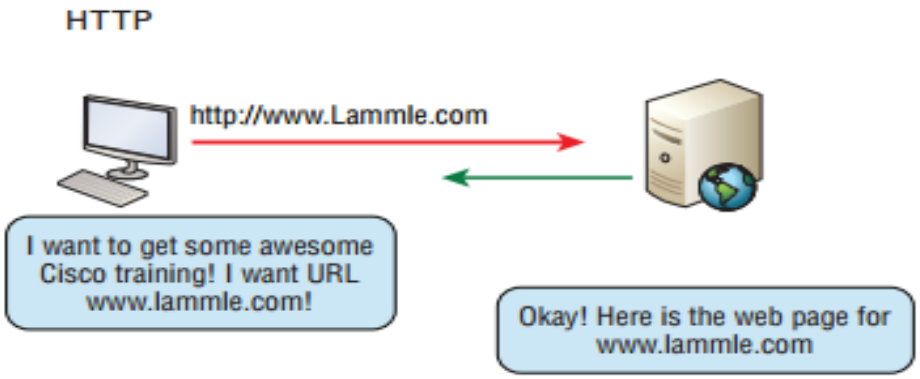
In addition, network devices can inform the NMS about problems as they occur so the network administrator is alerted.



Hypertext Transfer Protocol (HTTP)

It's used to manage communications between web browsers and web servers and opens the right resource when you click a link, wherever that resource may actually reside. In order for a browser to display a web page, it must find the exact server that has the right web page, plus the exact details that identify the information requested. The browser can understand what you need when you enter a Uniform Resource Locator (URL), which we usually refer to as a web address, e.g. http://www.lammle.com/forum and http://www.lammle.com/blog.

Each URL defines the protocol used to transfer data, the name of the server, and the particular web page on that server.



Hypertext Transfer Protocol Secure (HTTPS)

Hypertext Transfer Protocol Secure (HTTPS) is also known as Secure Hypertext Transfer Protocol. It uses Secure Sockets Layer (SSL). Sometimes you'll see it referred to as SHTTP or S-HTTP, which were slightly different protocols, but since Microsoft supported HTTPS, it became the de facto standard for securing web communication. But no matter-as indicated, it's a secure version of HTTP that arms you with a whole bunch of security tools for keeping transactions between a web browser and a server secure.

Domain Name Service (DNS)

The Domain Name System (DNS) is a directory lookup service that provides a mapping between the name of a host on the Internet and its numerical address. Domain Name Service (DNS)- resolves hostnames- to IP addresses specifically, Internet names, such as www.au.edu.etBut you don't have to actually use DNS. You just type in the IP address of any device you want to communicate with and find the IP address of a URL by using the Ping program.

For example, >ping www.cisco.com will return the IP address resolved by DNS.

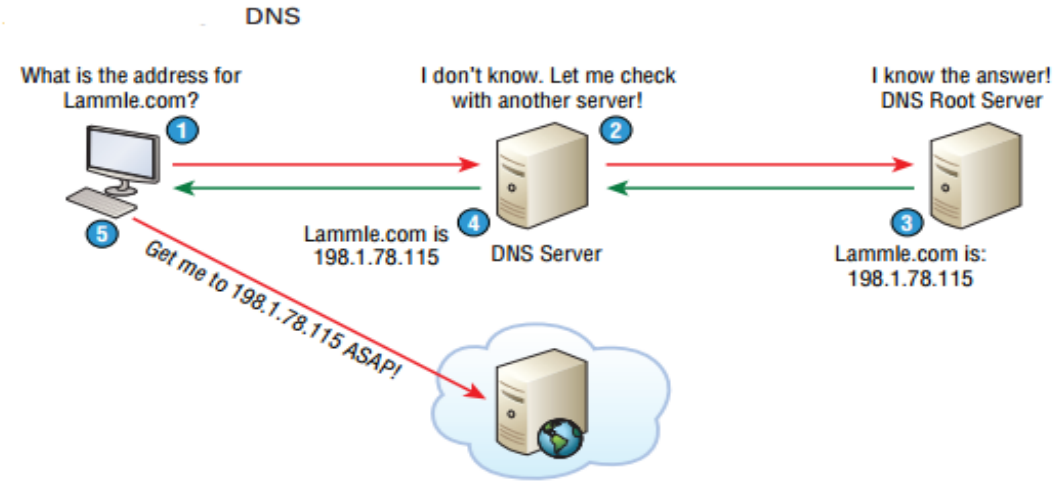
Domain Name System (DNS)

Resolves domain names to IP addresses and vice versa. An IP address identifies hosts on a network and the Internet as well, but DNS was designed to make our lives easier. The IP address would change and no one would know what the new one was. DNS allows you to use a domain name to specify an IP address.

Domain name

A domain name is represented by a series of character strings, called labels, separated by dots. Each label represents a level in the domain naming hierarchy. E.g In the domain name www.google.com, com is the top-level domain (TLD), google is the second-level domain, and www is the third-level domain. Each second-level domain can contain multiple third level domains. E.g In addition to www.google.com, Google also owns the following domains: news.google.com, maps.google.com, and mail.google.com. The very last section of the domain is called its top-level domain (TLD) name

Top-Level Domain	General Purpose	New TLDs	General Purpose
.com	U.S. Commercial	.biz	Business
.net	Network	.info	Information
.org	Nonprofit organization	.pro	Professional
.edu	U.S. Educational	.museum	Museums
.int	International	.aero	Aerospace industry
.mil	U.S. Military	.coop	Cooperative
.gov	U.S. Government		



Dynamic Host Configuration Protocol (DHCP)

Dynamic Host Configuration Protocol (DHCP) assigns IP addresses to hosts dynamically. It allows for easier administration and works well in small to very large network environments. Many types of hardware can be used as a DHCP server, including a Cisco router.

A DHCP address conflict occurs when two hosts use the same IP address. This sounds bad, and it is! A lot of information a DHCP server can provide to a host when the host is requesting an IP address from the DHCP server. Here's a list of the most common types of information a DHCP server can provide:

- IP address
- Subnet mask
- Domain name
- Default gateway (routers)
- DNS server address

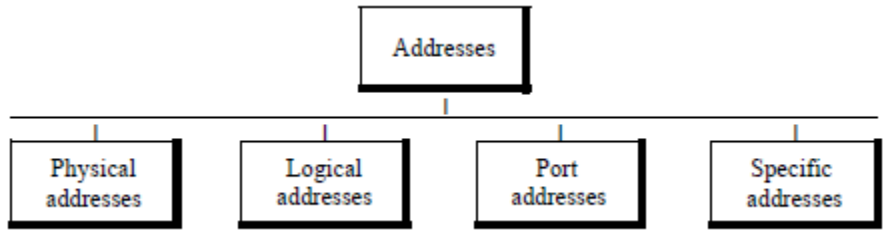
This is the four-step process a client takes to receive an IP address from a DHCP server:

1. The DHCP client broadcasts a DHCP Discover message looking for a DHCP server (Port 67).
2. The DHCP server that received the DHCP Discover message sends a layer 2 unicast DHCP Offer message back to the host.
3. The client then broadcasts to the server a DHCP Request message asking for the offered IP address and possibly other information.
4. The server finalizes the exchange with a unicast DHCP Acknowledgment message. Etc….

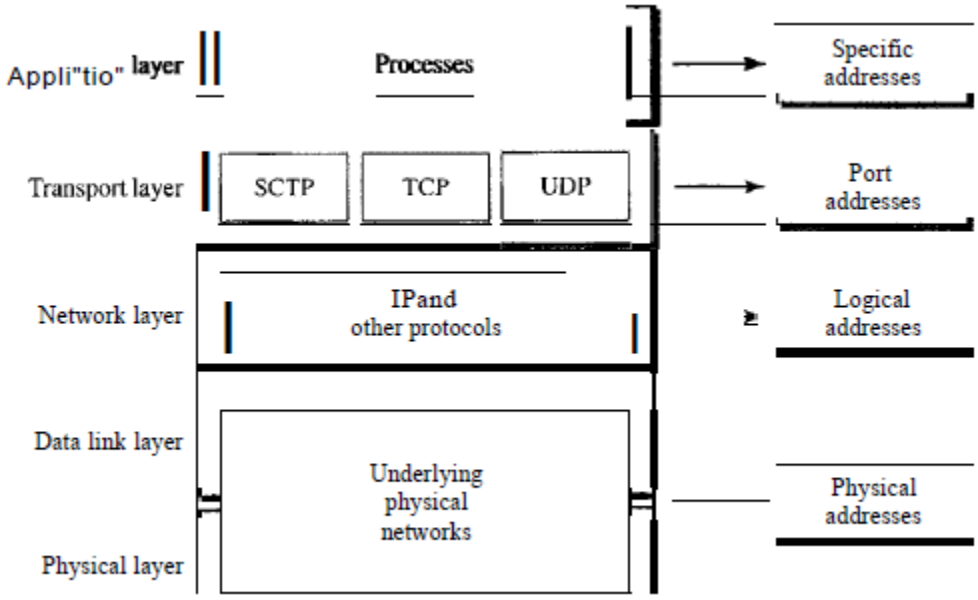
CHAPTER FIVE

INTERNET ADDRESSING

Four levels of addresses are used in an internet employing the *TCP/IP* protocols: *physical (link) addresses, logical (IP) addresses, port addresses, and specific addresses*



Each address is related to a specific layer in the TCPIIP architecture



Physical Addresses

The physical address, also known as the link address, is the address of a node as defined by its LAN or WAN. It is included in the frame used by the data link layer. It is the lowest-level address.

The physical addresses have authority over the network (LAN or WAN). The size and format of these addresses vary depending on the network. For example, Ethernet uses a 6-byte (48-bit) physical address that is imprinted on the network interface card (NIC).

Logical Addresses

Logical addresses are necessary for universal communications that are independent of underlying physical networks. Physical addresses are not adequate in an internetwork environment where different networks can have different address formats. A universal addressing system is needed in which each host can be identified uniquely, regardless of the underlying physical network.
The logical addresses are designed for this purpose. A logical address in the Internet is currently a 32-bit address that can uniquely define a host connected to the Internet. No two publicly addressed and visible hosts on the Internet can have the same IP address.

The physical addresses will change from hop to hop, but the logical addresses usually remain the same.

Port Addresses

The IP address and the physical address are necessary for a quantity of data to travel from a source to the destination host. However, arrival at the destination host is not the final objective of data communications on the Internet. A system that sends nothing but data from one computer to another is not complete. Today, computers are devices that can run multiple processes at the same time. The end objective of Internet communication is a process communicating with another process. For example, computer A can communicate with computer C by using TELNET. At the same time, computer A communicates with computer B by using the File Transfer Protocol (FTP). For these processes to receive data simultaneously, we need a method to label the different processes. In other words, they need addresses. In the TCPIIP architecture, the label assigned to a process is called a port address. A port address in TCPIIP is 16 bits in length.

The physical addresses change from hop to hop, but the logical and port addresses usually remain the same.

Specific Addresses

Some applications have user-friendly addresses that are designed for that specific address. Examples include the e-mail address (for example, forouzan@fhda.edu) and the Universal Resource Locator (URL) (for example, www.mhhe.com). The first defines the recipient of an e-mail; the second is used to find a document on the World Wide Web. These addresses, however, get changed to the corresponding port and logical addresses by the sending computer.

IP Addresses

1. IPv4

An **IPv4** address is a 32-bit address that *uniquely* and *universally* defines the connection of a device (for example, a computer or a router) to the Internet.

An IPv4 address is 32 bits long.

IPv4 addresses are unique. They are unique in the sense that each address defines one, and only one, connection to the Internet. Two devices on the Internet can never have the same address at the same time. We will see later that, by using some strategies, an address may be assigned to a device for a time period and then taken away and assigned to another device.

On the other hand, if a device operating at the network layer has *m* connections to the Internet, it needs to have *m* addresses. We will see later that a router is such a device.
The IPv4 addresses are universal in the sense that the addressing system must be accepted by any host that wants to be connected to the Internet.

Address Space
A protocol such as IPv4 that defines addresses has an address space. An address space is the total number of addresses used by the protocol. If a protocol uses *N* bits to define an address, the address space is *2N* because each bit can have two different values (0 or 1) and *N* bits can have 2^N values.
IPv4 uses 32-bit addresses, which means that the address space is 232 or 4,294,967,296 (more than 4 billion). This means that, theoretically, if there were no restrictions, more than 4 billion devices could be connected to the Internet.

Notations
There are two prevalent notations to show an IPv4 address: binary notation and dotted decimal notation.
Binary Notation
In binary notation, the IPv4 address is displayed as 32 bits. Each octet is often referred to as a byte. So it is common to hear an IPv4 address referred to as a 32-bit address or a 4-byte address. The following is an example of an IPv4 address in binary notation:

01110101 10010101 00011101 00000010

Dotted-Decimal Notation
To make the IPv4 address more compact and easier to read, Internet addresses are usually written in decimal form with a decimal point (dot) separating the bytes. The following is the **dotted~decimal** notation of the above address:

117.149.29.2
An IPv4 address in both binary and dotted-decimal notation. Note that because each byte (octet) is 8 bits, each number in dotted-decimal notation is a value ranging from 0 to 255.

Example 1
Change the following IPv4 addresses from binary notation to dotted-decimal notation.
a. 10000001 00001011 00001011 11101111

b. 11000001 10000011 00011011 11111111

Solution

We replace each group of 8 bits with its equivalent decimal number (see Appendix B) and add dots for separation.

a. 129.11.11.239

b. 193.131.27.255

Example 2

Change the following IPv4 addresses from dotted-decimal notation to binary notation.

a. 111.56.45.78

b. 221.34.7.82

Solution

We replace each decimal number with its binary equivalent

a.• 01101111 00111000 00101101 01001110

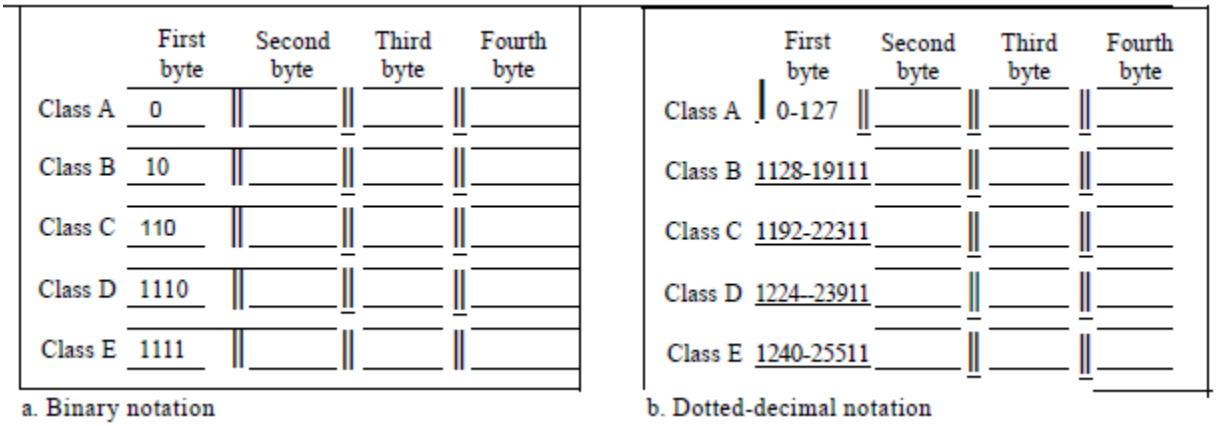
b. 11011101 00100010 00000111 01010010

Classful Addressing

IPv4 addressing, at its inception, used the concept of classes. This architecture is called classful addressing. Although this scheme is becoming obsolete, we briefly discuss it here to show the rationale behind classless addressing. In classful addressing, the address space is divided into five classes: A, B, C, D, and E. Each class occupies some part of the address space.

In classful addressing, the address space is divided into five classes: A, B, C, D, and E.

We can find the class of an address when given the address in binary notation or dotted-decimal notation. If the address is given in binary notation, the first few bits can immediately tell us the class of the address. If the address is given in decimal-dotted notation, the first byte defines the class. Both methods are shown in Figure below.



a. Binary notation

b. Dotted-decimal notation

Classes and Blocks

One problem with classful addressing is that each class is divided into a fixed number of blocks with each block having a fixed size as shown in Table

Class	*Number ofBlocks*	*Block Size*	*Application*
A	128	16,777,216	Unicast
B	16,384	65,536	Unicast
C	2,097,152	256	Unicast
D	1	268,435,456	Multicast
E	1	268,435,456	Reserved

In c1assful addressing, a large part of the available addresses were wasted.

Netid (Network Id) and Hostid

In classful addressing, an IP address in class A, B, or C is divided into netid and hostid. These parts are of varying lengths, depending on the class of the address. Note that the concept does not apply to classes D and E.

In class A, one byte defines the netid and three bytes define the hostid. In class B, two bytes define the netid and two bytes define the hostid. In class C, three bytes define the netid and one byte defines the hostid.

Mask

Although the length of the netid and hostid (in bits) is predetermined in classful addressing, we can also use a mask (also called the default mask), a 32-bit number made of contiguous Is followed by contiguous as. The masks for classes A, B, and C are shown in Table 19.2. The concept does not apply to classes D and E.

Table below *Default masks for classful addressing*

Class	*Binary*	*Dotted-Decimal*	*CIDR*
A	11111111 00000000 00000000 00000000	255.0.0.0	*18*
B	11111111 11111111 00000000 00000000	255.255.0.0	*116*
C	11111111 11111111 11111111 00000000	255.255.255.0	*124*

The mask can help us to find the netid and the hostid. For example, the mask for a class A address has eight 1s, which means the first 8 bits of any address in class A define the netid; the next 24 bits define the hostid.

The last column of Table 19.2 shows the mask in the form *In* where *n* can be 8, 16, or 24 in classful addressing. This notation is also called slash notation or Classless Interdomain Routing (CIDR) notation. The notation is used in classless addressing, which we will discuss later. We introduce it here because it can also be applied to classful addressing. We will show later that classful addressing is a special case of classless addressing.

Subnetting

During the era of classful addressing, subnetting was introduced. If an organization was granted a large block in class A or B, it could divide the addresses into several contiguous groups and assign each group to smaller networks (called subnets) or, in rare cases, share part of the addresses with neighbors. Subnetting increases the number of Is in the mask, as we will see later when we discuss classless addressing.

Classless Addressing

To overcome address depletion and give more organizations access to the Internet, classless addressing was designed and implemented. In this scheme, there are no classes, but the addresses are still granted in blocks.

Address Blocks

In classless addressing, when an entity, small or large, needs to be connected to the Internet, it is granted a block (range) of addresses. The size of the block (the number of addresses) varies based on the nature and size of the entity. For example, a household may be given only two addresses; a large organization may be given thousands of addresses. An ISP, as the Internet service provider, may be given thousands or hundreds of thousands based on the number of customers it may serve. Restriction to simplify the handling of addresses, the Internet authorities impose three restrictions on classless address blocks:

1. The addresses in a block must be contiguous, one after another.
2. The number of addresses in a block must be a power of 2 (I, 2, 4, 8, ...).
3. The first address must be evenly divisible by the number of addresses.

CHAPTER SIX
SUBNETTING

The need for sub netting

Classes of IP addresses offer a range from 256 to 16.8 million hosts. Subnetting separates a network into multiple logically defined segments, or subnets. To efficiently manage a limited supply of IP addresses, all classes can be subdivided into smaller sub networks or subnets. This process is known as subnetting.

The sub - netting process

To create the sub network structure, host bits must be reassigned as network bits which is often referred to as borrowing bits. The **starting point** for this process is always the leftmost bit of the host. That is he one closest to the last network octet.

Subnet addresses include:

- The Class A, Class B, and Class C network portion,
- a subnet field and
- a host field.

The subnet field and the host field are created from the original host portion of the major IP address. This is done by assigning bits from the host portion to the original network portion of the address. Subnets have sub network ID (subnet ID) just as networks have network IDs. Subnet IDs are found by replacing all host fields with 0s.

Sub netting Advantages

The subnet field and the host field are created from the original host portion of the major IP address. This is done by assigning bits from the host portion to the original network portion of the address. Subnets have sub network ID (subnet ID) just as networks have network IDs. Subnet IDs are found by replacing all host fields with 0s.

Borrowing a bits

To determine the number of bits to be used, the network designer needs to calculate how many hosts the largest sub network requires and the number of sub networks needed Large number of subnets means fewer hosts and a large number of hosts means fewer subnets Total number of subnets is 2^bits borrowed Total number of hosts is 2^remaining host bits Example if three bits are borrowed from a class C address, total number of subnets is 8 (2^3) and total number of hosts is 32 (2^5)

Positional value of bits

Bits borrowed	1	2	3	4	5	6	7	8
Value	128	64	32	16	8	4	2	1

Value is the position value of the bits borrowed.

Example

Q - What is the value of 01010110 in decimal?

A - 0 + 64 + 0 + 16 + 0 + 4 + 2 + 0 = 86

Usable subnets & Usable Hosts

Among the available subnets, it is not advised to use the following two subnets:

- The subnet with all 0's in the subnet field
- The subnet with all 1's in the subnet field

If subnet zero (all 0's in the subnet field) is used, it means that a **network** and a **subnet** have the same address. If the last subnet (all 1's in the subnet field) is used, it means that the **network broadcast address** and a **subnet** have the same address. Hence usable subnets will be **2^bits borrowed – 2**. Example if three bits are borrowed from a class C address, total number of *usable* subnets is 6 (2^3 - 2) and total number of usable hosts is 30 (2^5 - 2)

Subnet Masks

For the subnet address scheme to work, every machine on the network must know which part of the host address will be used as the subnet address. This is accomplished by assigning a subnet mask to each machine. A **subnet mask is a 32-bit** value that allows the recipient of IP packets to distinguish the network portion of the IP address from the host portion of the IP address

A subnet mask is composed of 1s and 0s where:

- The 1s in the subnet mask represent the positions that refer to the network or subnet addresses
- The 0s in the subnet mask represent the positions that refer to the host address

Default subnet masks

Not all networks need subnets, meaning they use the default subnet mask. This is basically the same as saying that a network doesn't have a subnet address. Here is default subnet mask for Classes A, B, and C

- Class A - **network.node.node.node** Subnet mask: 255.0.0.0
- Class B **network.network.node.node** Subnet mask:255.255.0.0
- Class C - **network. network.network.node** Subnet mask: 255.255.255.0

These default subnet masks show the minimum number of 1's you can have in a subnet mask for each class.

Specifying subnets

- Example if three bits are borrowed from a class C address, the subnet mask is 255.255.255.224
- Subnets may also be represented, in a slash format.
- For example, /24 indicates that the total bits that were used for the network and sub network portion is 24
- The subnet mask 255.255.255.224 in slash format is /27. (224=11100000)

Number of bits borrowed from a class C address, positional value of each bit and resulting mask (in number and slash format).

Slash format	/25	/26	/27	/28	/29	/30	N/A	N/A
Mask	128	192	224	240	248	252	254	255
Bits borrowed	1	2	3	4	5	6	7	8
Value	128	64	32	16	8	4	2	1

Sub netting Class C addresses

Example 1

Let us subnet the network address 192.168.10.0 with a subnet mask 255.255.255.192 or in slash format /26

- (192 is 11000000)

Q - How many usable subnets do we have?

A - Since 192 is 2 bits on (**11**000000), the answer would be 2^2 – 2 = 2

Q - How many usable hosts per subnet do we have?

A - We have 6 host bits off (11**000000**), so the answer would be 2^6 – 2 = 62 hosts

Q - What are the subnet IDs?

A - We vary the borrowed bits (00, 01, 10, 11).

So the subnets are 192.168.10.0, 192.168.10.64, 192.168.10.128, 192.168.10.192

Q - What are the valid or usable subnets.

A - The ones which do not have all 0's or all 1's in the subnet field, namely 192.168.10.64 and 192.168.10.128

Q - What's the broadcast address for the valid subnets?

A - The valid subnets start with 01 and 10. The broadcast address for these two addresses will have 01111111 and 10111111. Which are 127 and 191. So the broadcast addresses will be 192.168.10.127 and 192.168.10.191. As a shortcut you can follow this rule: The number right before the value of the next subnet is all host bits turned on and equals the broadcast address.

Q - What are the valid hosts?

A - These are the numbers between the subnet ID and broadcast address

The hosts for the first valid subnet are:

- 192.168.10.65, 192.168.10.66, …, 192.168.10.126

The hosts for the second valid subnet are:

- 192.168.10. 129, 192.168.10.130, …, 192.168.10.190

Example 2

Now let us subnet the network address 192.168.10.0, this time with a subnet mask 255.255.255.224 or in slash format /27

Q - How many subnets do we have?

A - Since 224 is 3 bits on (**111**00000), the answer would be 2^3 – 2 = 6

Q - How many hosts per subnet do we have?

A - We have 6 host bits off (111**00000**), so the answer would be 2^5 – 2 = 30 hosts

Q - What are the subnet IDs?

A We vary the borrowed bits (000, 001, 010, 011, 100, 101, 110, 111). So the subnets are 192.168.10.0, 192.168.10.32, 192.168.10.64, 192.168.10.96, 192.168.10.128, 192.168.10.160, 192.168.10.192, 192.168.10.224

Q - What are the valid or usable subnets?

A - 192.168.10.32, 192.168.10.64, 192.*1*68.10.96, 192.168.10.128, 192.168.10.160, 192.168.10.192

Q - What's the broadcast address for the valid subnets?

A - The number right before the value of the next subnet is all host bits turned on and equals the broadcast address – 192.168.10.63, 192.168.10.95, 192.168.10.127, 192.168.10.159, 192.168.10.191, 192.168.10.223

Q - What are the valid hosts?

192.168.10.33 – 192.168.10.62	**192.168.10.129 – 192.168.10.161**
192.168.10.65 – 192.168.10.94	**192.168.10.161 – 192.168.10.193**
192.168.10.97 – 192.168.10.128	**192.168.10.193– 192.168.10.222**

Example 3

Subnet the network address 192.168.10.0, with a subnet mask 255.255.255.248 (/28)

Q - How many subnets do we have?

A - Since 248 is 4 bits on (**1111**0000), 2^4 – 2 = 14

Q - How many hosts per subnet do we have?

A - We have 6 host bits off (1111**0000**), 2^4 – 2 = 14

Q - What are the subnet IDs?

A - We vary the borrowed bits (0000, 0001, 0010, 0011, 0100, 0101, 0110, 0111, 1000, 1001, 1010, 1011, 1100, 1101, 1110, 1111). So the subnets ID's are:

192.168.10.0, 192.168.10.16, 192.168.10.32, 192.168.10.48, 192.168.10.64, 192.168.10.80, 192.168.10.96, 192.168.10.112, 192.168.10.128, 192.168.10.144, 192.168.10.160, 192.168.10.176, 192.168.10.192, 192.168.10.208, 192.168.10.224**,** **192.168.10.240**

Q - What are the valid or usable subnets?

Q - What's the broadcast address for the valid subnets?

A - 192.168.10.31, 192.168.10.47, 192.168.10.63, 192.168.10.79, 192.168.10.95, 192.168.10.111, 192.168.10.127, 192.168.10.143, 192.168.10.159, 192.168.10.175, 192.168.10.191, 192.168.10.107, 192.168.10.223, 192.168.10.239

Q - What are the valid hosts?

192.168.10.17 – 192.168.10.30, 192.168.10.33 – 192.168.10.46, 192.168.10.49 – 192.168.10.62, 192.168.10.65 – 192.168.10.78, 192.168.10.81 – 192.168.10.94, 192.168.10.97 – 192.168.10.110, 192.168.10.113 –192.168.10.126, 192.168.10.129 – 192.168.10.142, 192.168.10.145 – 192.168.10.158, 192.168.10.161 – 192.168.10.174192.168.10.177 –

192.168.10.190, , 192.168.10.193 – 192.168.10.106, 192.168.10.109 – 192.168.10.222, 192.168.10.225 – 192.168.10.238.

Subnetwork #	Subnetwork ID	Host Range	Broadcast ID
0	192.168.10.0	.1--.30	192.168.10.31
1	192.168.10.32	.33--.62	192.168.10.63
2	192.168.10.64	.65--.94	192.168.10.95
3	192.168.10.96	.97--.126	192.168.10.127
4	192.168.10.128	.129--.158	192.168.10.159
5	192.168.10.160	.161--.190	192.168.10.191
6	192.168.10.192	.193--.222	192.168.10.223
7	192.168.10.224	.225--.254	192.168.10.255

Calculating class A and B networks

The Class A and B sub netting procedure is identical to the process for Class C, except there may be significantly more bits involved.

Assigning 12 bits of a Class B address to the subnet field creates a subnet mask of 255.255.255.240 or /28.

All eight bits were assigned in the third octet resulting in 255, the total value of all eight bits. Four bits were assigned in the fourth octet resulting in 240.

Possible Class B subnet masks

255.255.128.0 (/17)
255.255.192.0 (/18)
255.255.224.0 (/19)
255.255.240.0 (/20)
255.255.248.0 (/21)
255.255.252.0 (/22)
255.255.254.0 (/23)
255.255.255.0 (/24)
255.255.255.128 (/25)
255.255.255.192 (/26)
255.255.255.224 (/27)
255.255.255.240 (/28)
255.255.255.248 (/29)
255.255.255.252 (/30)

Sub netting Class B addresses

Example 1

172.16.0.0 = Network address

255.255.192.0 = Subnet mask

Q - How many Subnets?

A - 2^2 – 2 = 2.

Q - How many Hosts per subnet?

2^14 – 2 = 16,382. (6 bits in the third octet, and 8 in the fourth)

Q - Subnet IDs of valid subnets?

A - 172.16.64.0 and 172.16.128.0

Q Broadcast address for each subnet and valid hosts?

A Below is the two subnets available and the address of each:

Subnet	172.16.64.0	172.16.128.0
First host	172.16.64.1	172.16.128.1
Last host	172.16.127.254	172.16.191.254
Broadcast	172.16.127.255	172.16.191.255

Example 2

172.16.0.0 = Network address

255.255.240.0 = Subnet mask

Q How many Subnets?

A 2^4 – 2 = 14

Q How many Hosts per subnet?

2^12 – 2 = 4094

Q Subnet IDs of valid subnets?

A 172.16.16.0 and 172.16.32.0, …, 172.16.224.0

Q Broadcast address for each subnet and valid hosts?

A Below is the subnets available and the address of each:

Subnet	172.16.16.0	172.16.32.0	**…**
First host	172.16.16.1	172.16.32.1	**…**
Last host	172.16.31.254	172.16.47.254	**…**
Broadcast	172.16.31.255	172.16.47.255	**…**

Possible Class A subnet masks

255.128.0.0 (/9)	255.255.240.0 (/20)
255.192.0.0 (/10)	255.255.248.0 (/21)
255.224.0.0 (/11)	255.255.252.0 (/22)
255.240.0.0 (/12)	255.255.254.0 (/23)
255.248.0.0 (/13)	255.255.255.0 (/24)
255.252.0.0 (/14)	255.255.255.128 (/25)
255.254.0.0 (/15)	255.255.255.192 (/26)
255.255.0.0 (/16)	255.255.255.224 (/27)
255.255.128.0 (/17)	255.255.255.240 (/28)
255.255.192.0 (/18)	255.255.255.248 (/29)
255.255.224.0 (/19)	255.255.255.252 (/30)

Sub netting Class A addresses

Example 1

10.0.0.0 = Network address

255.255.0.0 (/16) = Subnet mask

Q Subnets?

A 2^8 – 2 = 254

Q Hosts?

A 2^16 – 2 = 65,534

Q Valid subnets?

A 10.1.0.0, 10.2.0.0, 10.3.0.0, …, 10.254.0.0

Q Broadcast address for each subnet and valid hosts?

Subnet	10.1.0.0	… 10.254.0.0
First host	10.1.0.1	… 10.254.0.1
Last host	10.1.255.254	… 10.254.255.254
Broadcast	10.1.255.255	… 10.254.255.255

Example 2

10.0.0.0 = Network address

255.255.240.0 (/20) = Subnet mask

Q Subnets?

2^12 – 2 = 4094

Q Hosts?

A 2^12 – 2 = 4094

Q Valid subnets?

A

Subnet	10.1.0.0, 10.1.16.0,**…,** 10.255.224.0
First host	10.1.0.1, 10.1.16.1,**…,** 10.255.224.1
Last host	10.1.15.254, 10.1.31.254,**…**, 10.255.239.254
Broadcast	10.1.15.255, 10.1.31.255,**…**, 10.255.239.255

ANDing

The subnet mask gives routers the information required to determine in which network and subnet a particular host resides.

Routers make an AND operation between the subnet mask and the destination address (ANDing) to determine the subnet ID of the destination address. This information is required for routing purposes.

Packet address	201.10.11.65	11001001.00001010.00001011.01000001
AND		
Mask	255.255.255.224	11111111.11111111.11111111.11100000
Subnetwork ID	201.10.11.64	11001001.00001010.00001011.01000000

Example

Address of host X is 192.168.54.84

Subnet mask of host X is 255.255.255.224

The sub network ID for host X: 192.168.54.84 AND 255.255.255.224 = **192.168.54.64** is the network address for host X

CHAPTER SEVEN

CONNECTING DEVICES (LAN and WAN Technologies)

Repeaters

A repeater is a device that operates only in the physical layer. Signals that carry information within a network can travel a fixed distance before attenuation endangers the integrity of the data. A repeater receives a signal and, before it becomes too weak or corrupted, regenerates the original bit pattern. The repeater then sends the refreshed signal. A repeater can extend the physical length of a LAN.

A repeater does not actually connect two LANs; it connects two segments of the same LAN. The segments connected are still part of one single LAN. A repeater is not a device that can connect two LANs of different protocols.

Hubs

Passive Hubs

A passive hub is just a connector. It connects the wires coming from different branches. In a star topology Ethernet LAN, a passive hub is just a point where the signals coming from different stations collide; the hub is the collision point. This type of a hub is part of the media; its location in the Internet model is below the physical layer.

Active Hubs

An active hub is actually a multipart repeater. It is normally used to create connections between stations in a physical star topology.

Bridges

A bridge operates in both the physical and the data link layer. As a physical layer device, it regenerates the signal it receives. As a data link layer device, the bridge can check the physical (MAC) addresses (source and destination) contained in the frame.

Two-Layer Switches

When we use the term *switch,* we must be careful because a switch can mean two different things. We must clarify the term by adding the level at which the device operates. We can have a two-layer switch or a three-layer switch. A **three-layer switch** is used at the network layer; it is a kind of router. The **two-layer switch** performs at the physical and data link layers.

A two-layer switch is a bridge, a bridge with many ports and a design that allows better (faster) performance. A bridge with a few ports can connect a few LANs together. A bridge with many ports may be able to allocate a unique port to each station, with each station on its own independent entity. This means no competing traffic (no collision, as we saw in Ethernet).

A two-layer switch, as a bridge does, makes a filtering decision based on the MAC address of the frame it received. However, a two-layer switch can be more sophisticated. It can have a buffer to hold the frames for processing. It can have a switching factor that forwards the frames faster. Some new two-layer switches, called *cut-through* switches, have been designed to forward the frame as soon as they check the MAC addresses in the header of the frame.

Routers

A router is a three-layer device that routes packets based on their logical addresses (host-to-host addressing). A router normally connects LANs and WANs in the Internet and has a routing table that is used for making decisions about the route. The routing tables are normally dynamic and are updated using routing protocols.

Three-Layer Switches

A three-layer switch is a router, but a faster and more sophisticated. The switching fabric in a three-layer switch allows faster table lookup and forwarding. In this book, we use the terms *router* and *three-layer switch* interchangeably.

Gateway

Although some textbooks use the terms *gateway* and *router* interchangeably, most of the literature distinguishes between the two. A gateway is normally a computer that operates in all five layers of the Internet or seven layers of OSI model. A gateway takes an application message, reads it, and interprets it. This means that it can be used as a connecting device between two internetworks that use different models. For example, a network designed to use the OSI model can be connected to another network using the Internet model. The gateway connecting the two systems can take a frame as it arrives from the first system, move it up to the OSI application layer, and remove the message. Gateways can provide security.

CHAPTER EIGHT
COMPUTER NETWORK SECURITY BASICS

During initial days of internet, its use was limited to military and universities for research and development purpose. Later when all networks merged together and formed internet, the data used to travel through public transit network. Common people may send the data that can be highly sensitive such as their bank credentials, username and passwords, personal documents, online shopping details, or confidential documents.

All security threats are intentional i.e. they occur only if intentionally triggered. Security threats can be divided into the following categories:

Interruption

Interruption is a security threat in which availability of resources is attacked. For example, a user is unable to access its web-server or the web-server is hijacked.

Privacy-Breach

In this threat, the privacy of a user is compromised. Someone, who is not the authorized person is accessing or intercepting data sent or received by the original authenticated user.

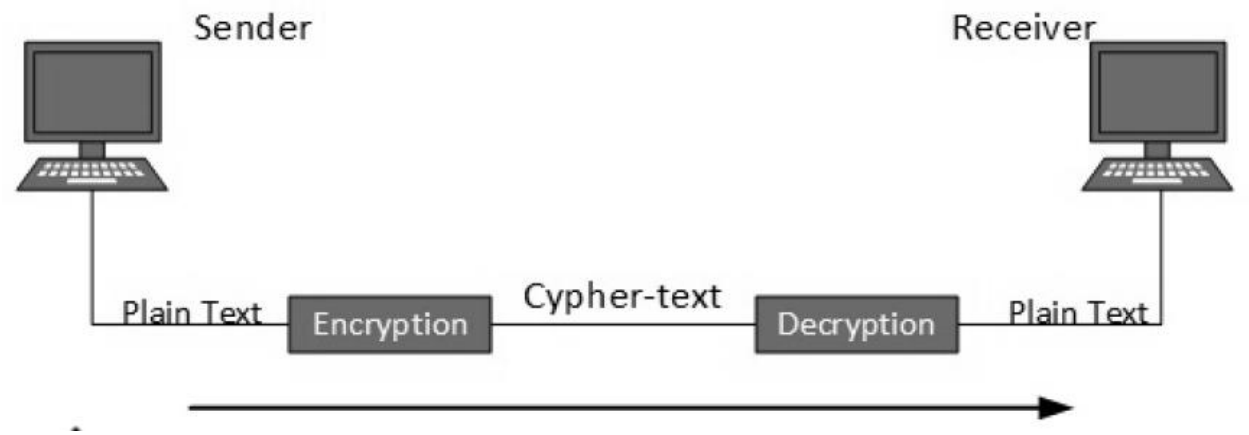
Integrity

This type of threat includes any alteration or modification in the original context of communication. The attacker intercepts and receives the data sent by the sender and the attacker then either modifies or generates false data and sends to the receiver. The receiver receives the data assuming that it is being sent by the original Sender.

Authenticity

This threat occurs when an attacker or a security violator poses as a genuine person and accesses the resources or communicates with other genuine users.

No technique in the present world can provide 100% security. But steps can be taken to secure data while it travels in unsecured network or internet. The most widely used technique is Cryptography.



Cryptography is a technique to encrypt the plain-text data which makes it difficult to understand and interpret. There are several cryptographic algorithms available present day as described below:

- Secret Key
- Public Key
- Message Digest

Secret Key Encryption

Both sender and receiver have one secret key. This secret key is used to encrypt the data at sender's end. After the data is encrypted, it is sent on the public domain to the receiver. Because the receiver knows and has the Secret Key, the encrypted data packets can easily be decrypted.

Example of secret key encryption is Data Encryption Standard (DES). In Secret Key encryption, it is required to have a separate key for each host on the network making it difficult to manage.

Public Key Encryption

In this encryption system, every user has its own Secret Key and it is not in the shared domain. The secret key is never revealed on public domain. Along with secret key, every user has its own but public key. Public key is always made public and is used by Senders to encrypt the data. When the user receives the encrypted data, he can easily decrypt it by using its own Secret Key.

Example of public key encryption is Rivest-Shamir-Adleman (RSA)..

Message Digest

In this method, actual data is not sent; instead a hash value is calculated and sent. The other end user, computes its own hash value and compares with the one just received. If both hash values are matched, then it is accepted; otherwise rejected.

Example of Message Digest is MD5 hashing. It is mostly used in authentication where user password is cross checked with the one saved on the server.