DATA COMMUNICATON AND COMPUTER NETWORK

1.1 DATA COMMUNICATION

In General Term, communication means sharing of Information / ideas. But in networking, communication means sharing of Data/Information. Data refers to facts, concepts, instructions etc. It may be Text, Picture, Audio, Video, etc. In Computer system, data are represented by 0 or 1 (Binary form).

Data Communication is the method to exchange of data in the form of 0 or 1 between two / more devices by following some rules (protocols) through some transmission media. The transmission may be Guided or Unguided.

Characteristics

For Data Communication, we need : a) Hardwarer and b) Software.
The effectiveness of data communication depends on :

- Delivery
- Accuracy
- Timeliness

Delivery - The system must deliver data to the correct destination. Data must be received by intended device/user

Accuracy - The system must deliver data accurately. Data that have been altered in transmission and delivered incorrectly are not usable

Timeliness - The system must deliver data in time. Late delivered data are sometimes useless

Components required for communication

- **Sender** - Sender is a device which sends data/message.
- **Receiver** - Receiver is a device which receives data/message
- **Message** - Message is the data/info^{n} that is to be transmitted.
- **Transmission Medium**-Medium is the physical path by which a message can be sent. It may be guided media or un-guided media
- **Protocol** - Protocol is a set of rules/agreement that govern/control the data communication. Without protocol, 2 devices may be connected but not communicated

1.2 NETWORKS

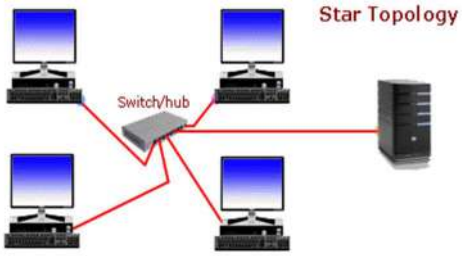
It is the network, where two or more devices/nodes connected with each other directly or indirectly for the purpose of Sharing information (like: analog/ digital) and resources. The resources may be software resources or hardware devices.

Types of Network

A Network can be of :

a) **Wire based Network**

When a group of computers/ devices connected with each other ***using cable/wire*** as a medium to transmit data, it is known as wired network.



Example : LAN in a building, LAN in a whole campus

b) **Wireless Network**

When a group of computers/ devices connected with each other ***without using cable/wire*** to transmit data, it is known as wireless network.

Example : Accessing internet through Mobile, Accessing internet through HotSpot

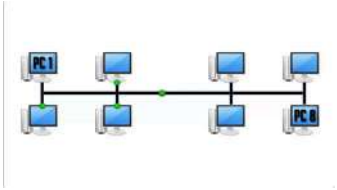
Categories of Network

Network can be categorized as :

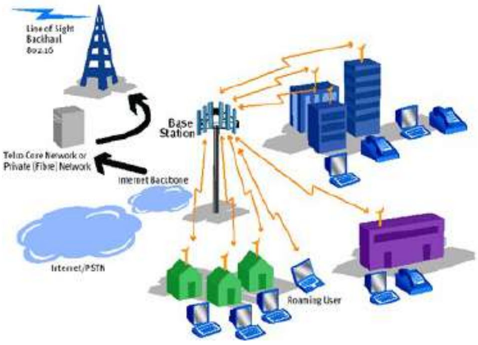
a) Local Area Network (LAN)
b) Metropolitan Area Network (MAN)
c) Wide Area Network(WAN)
d) Private Area Network(PAN)

a) **LAN** : A Local Area Network is a computer network that interconnects computers ***within a limited area***.

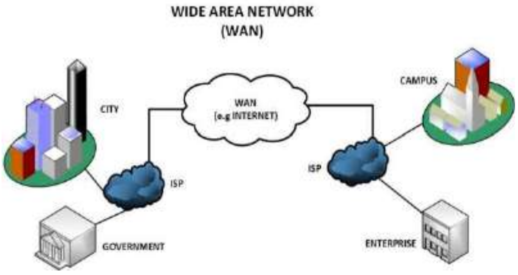
Example : a residence, school, laboratory, university campus or office building.



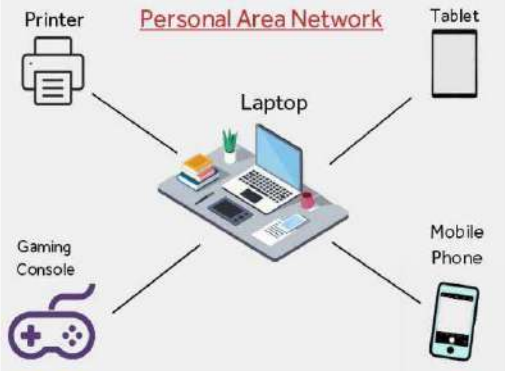
b) **MAN** : A Metropolitan Area Network (MAN) is a network that interconnects computers in a ***specific geographical area***. It is larger than the area covered by LAN but smaller than the area covered by WAN (***between 5 to 50 KM*** in range)
Example : Local cable operator provide services throughout the city


Line of Sight Backhaul 802.16
Base Station
Telco Core Network or Private (Fibre) Network
Internet Backbone
Internet/PSTN
Roaming User

c) **WAN** : A Wide Area Network (WAN) is a ***collection of LANs, MANs or other*** networks that communicate with one another. It is a telecomm network that ***extends over a large geographical area*** for the purpose of computer networking. A WAN is a ***network of networks.*** **Internet** the world's largest WAN.



d) **PAN** : A personal area network (PAN) is a computer network for interconnecting electronic devices on an individual person's workspace. A PAN provides data transmission among devices such as computers, smartphones, tablets, and personal digital assistants. A PAN may be wireless or carried over wired interfaces such as USB. A wireless personal area network (WPAN) is a PAN carried over a low-powered, short-distance wireless network technology such as IrDA, Wireless USB, Bluetooth. The reach of a WPAN varies from a few centimetres to a few meters.



1.3 PROTOCOL AND ARCHITECTURE, STANDARDS, OSI AND TCP/IP

Standards and **Protocols** are used to define representation and interaction modes within a network and to make certain functions available. Standards and protocols usually come in groups and work together and constitute **protocol suites** or protocol stacks. New standards and protocols are continuously being introduced. They fit within general frameworks called **Architectures.**

International Standards Organization (ISO), develop one standard architecture, called the **Open Systems Interconnect (OSI)**.

Another network architecture that is not as general as the OSI architecture, is the **TCP/IP Architecture**.

Both the OSI and the TCP/IP architectures are **layered architectures,** where a higher level layer uses the services provided by the layer immediately below it. Across a network communicating entities communicate exclusively at the same layer (i.e. if a sender entity at layer i sends a packet P, the receiver entity at layer i will receive P; these two are called **peer entities**).

NETWORK ARCHITECTURE

- **Network architecture** understood as the set of layers and layer protocols that constitute the communication system. Network architectures offer different ways of solving a critical issue when built a network: transfer data quickly and efficiently by the devices that make up the network.
- The type of network architecture used will not only determine the network topology but also define how network nodes access those media. There are different types of network architecture, all of them with their strategy to conduct information over the network.
- Network Architecture defines the communications products and services, which ensure that the various components can work together. In the early days of data communication systems, the majority of communications were between the DTE and the host computer.
- Recent computer systems link with other systems to form a network. Hence, the network architecture represents a systemization of the various kinds of protocols needed to build a network.

Types of Network Architecture

Ethernet

- Ethernet provides network access using multiple cover perception access with collision detection or CSMA / CD (carrier sense multiple access with collision detection). This network access strategy is basically that each component of the network or node listens before transmitting the information packets. If two nodes transmit at the same time, a collision occurs. When a collision is detected, the computer interrupts the transmission and waits for the line to be free.

Token Ring

- IBM Token Ring is a faster and safer network that uses the signal token as a strategy to access the communication channel. Token Ring networks connected in a star-shaped topology through a Multistation Access Unit (**MAU**) that provides the central connection for the nodes of the network. The ring through which the signal or token circulates (the token travels in only one direction) is a logical ring included within the MAU.

FDDI (Fiber Distributed Data Interface)

- The Fiber Distributed Data Interface (FDDI) is an architecture that provides a high speed and high capacity environment that can be used to connect several different types of networks. FDDI uses Fiber Optic Cables and configured in a ring topology. FDDI uses the signal or token pass as a method of access to the communication channel and can operate at high speeds (almost all implementations work at 100Mbps, but data can also transfer at higher speeds).

AppleTalk

- AppleTalk uses a unique addressing system to determine the address of the nodes included in the network. AppleTalk is the network architecture used by Apple Macintosh computers. The wiring system that allows Macintosh computers to connect is called local talk and uses twisted pair cables with a special adapter for Macintosh. When a Macintosh connected to the network is turned on, that computer generates a random address and transmits it over the network. This random address becomes your network address (provided that no other Macintosh computer uses that same address; otherwise, the computer has to continue generating random addresses until it finds one that has not used).

Types of Network Protocols

- **Network Protocols** are a set of rules governing exchange of information in an easy, reliable and secure way. Protocol means, a set of mutually accepted and implemented rules at both ends of the communications channel for the proper exchange of information. By adopting these rules, two devices can communicate with each other and can interchange information. Each protocol is defined in different terms and different use with unique name. Message travel from sender to receiver via a medium (The medium is the physical path over which a message travels) using a protocol.

- All data of *protocols* are stored in binary information. Protocol language is a mixture of bits, characters, integers, etc. Each of it has its own access method of exchanging data over a computer network, such as LAN, Internet, Intranet, etc. One of the most common and known protocol example is HTTP (used over the www). There are different protocols used in internet that are
  - ❖ TCP/IP (Transmission Control Protocol/Internet Protocol)
  - ❖ ARP (Address Resolution Protocol)
  - ❖ DHCP (Dynamic Host Configuration Protocol
  - ❖ DNS (Domain Name System)
  - ❖ FTP (File Transfer Protocol)
  - ❖ PPP (Point-to-point Protocol)

Network Models :

OSI Model :

OSI stands for **Open System Interconnection.** It is a reference model that describes how information from a s/w application in one computer moves through a physical medium to the s/w application in another computer. OSI consists of seven layers, and each layer performs a particular network function. OSI model was developed by the International Organization for Standardization (ISO) in 1984. It is now considered as an architectural model for the inter-computer communications.

OSI model divides the whole task into seven smaller and manageable tasks. Each layer is assigned a particular task. Each layer is self-contained, so that task assigned to each layer can be performed independently. It simply tells what each layer should do by defining its input and output data. It is up to network architects to implement the layers according to their needs and resources available.

These are the 07 layers of the OSI model −

- Physical Layer
- Data link Layer

- Network Layer
- Transport Layer
- Session Layer
- Presentation Layer
- Application Layer

Physical layer –

It is the first layer that physically connects the two systems that need to communicate. It transmits data in bits and manages simplex or duplex transmission by modem. It also manages Network Interface Card's hardware interface to the network, like cabling, cable terminators, topography, voltage levels, etc.

Data link layer –

It is the firmware layer of Network Interface Card. It assembles datagrams into frames and adds start and stop flags to each frame. It also resolves problems caused by damaged, lost or duplicate frames.

Network layer –

It is concerned with routing, switching and controlling flow of information between the workstations. It also breaks down transport layer datagrams into smaller datagrams.

Transport layer –

Till the session layer, file is in its own form. Transport layer breaks it down into data frames, provides error checking at network segment level and prevents a fast host from overrunning a slower one. Transport layer isolates the upper layers from network hardware.

Session layer –

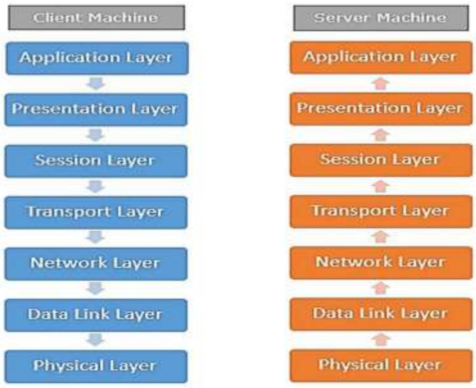
This layer is responsible for establishing a session between two workstations that want to exchange data.

Presentation layer –

This layer is concerned with correct representation of data, i.e. syntax and semantics of information. It controls file level security and is also responsible for converting data to network standards.

Application layer –

It is the topmost layer of the network that is responsible for sending application requests by the user to the lower levels. Typical applications include file transfer, E-mail, remote logon, data entry, etc.



TCP/IP Model :

TCP/IP stands for **Transmission Control Protocol/Internet Protocol**. TCP/IP is a set of layered protocols used for communication over the Internet. The communication model of this suite is client-server model. A computer that sends a request is the client and a computer to which the request is sent is the server. The TCP/IP model was developed prior to the OSI model. The TCP/IP model is not exactly similar to the OSI model.

The TCP/IP model consists of five layers. TCP/IP is a hierarchical protocol made up of interactive modules, and each of them provides specific functionality. The FIVE layers of this model are :

- Physical Layer
- Data link Layer
- Network Layer
- Transport Layer
- Application Layer

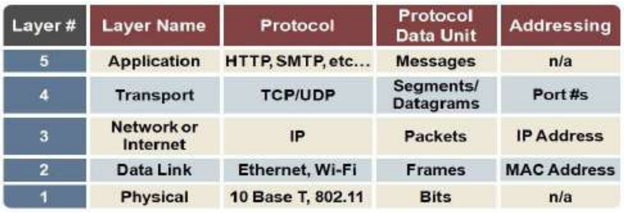
Application layer –
The Application layer is the layer, where applications requiring network communications live. Examples of these applications include email clients and web browsers. These applications use the Transport Layer to send requests to connect to remote hosts. Application layer protocols like HTTP and FTP are used.

Transport layer –
The Transport layer establishes the connection between applications running on different hosts. It uses TCP for reliable connections and UDP for fast connections. Data is transmitted in form of datagrams using the Transmission Control Protocol (TCP). TCP is responsible for breaking up data at the client side and then reassembling it on the server side. It keeps track of the processes running in the applications above it by assigning port numbers to them and uses the Network layer to access the TCP/IP network.

Network layer –
The Network layer is responsible for creating the packets that move across the network.
It uses IP addresses to identify the packet’s source and destination. Network layer connection is established using Internet Protocol (IP) at the network layer. Every machine connected to the Internet is assigned an address called IP address by the protocol to easily identify source and destination machines.

Data link layer –
Actual data transmission in bits occurs at the data link layer using the destination address provided by network layer. The Data Link layer is responsible for creating the frames that move across the network. These frames encapsulate the packets and use MAC addresses to identify the source and destination.

Physical Layer -
The Physical layer encodes and decodes the bits found in a frame and includes the transceiver that drives and receives the signals on the network.



Layer #	Layer Name	Protocol	Protocol Data Unit	Addressing
5	Application	HTTP, SMTP, etc...	Messages	n/a
4	Transport	TCP/UDP	Segments/ Datagrams	Port #s
3	Network or Internet	IP	Packets	IP Address
2	Data Link	Ethernet, Wi-Fi	Frames	MAC Address
1	Physical	10 Base T, 802.11	Bits	n/a

UNIT-2

DATA TRANSMISSION AND MEDIA

DATA TRANSMISSION CONCEPT & TERMINOLOGY :

DATA:

Data refers to information that conveys some meaning based on some agreed up rules between sender and receiver. It comes in a variety of forms like : text, audio, graphics, video, animation etc.

Data can be categorized into : a) analog data and b) digital data

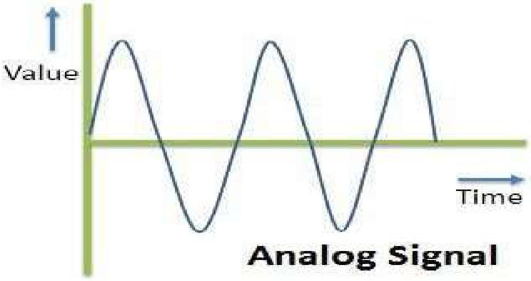
Analog data take on continuous values on some interval. Data that are collected from the real world with the help of transducers are continuous valued or analog in nature. Example :

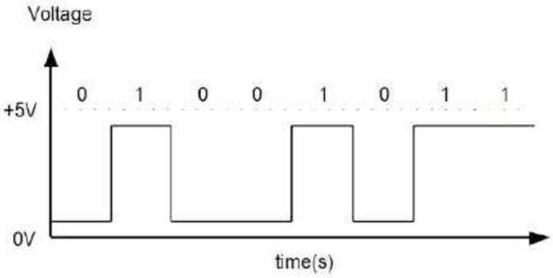
Digital data take on discrete values. Values are represented by a 7-bit code

SIGNAL :

It is a electrical, electronic or optical representation of data, which can be sent over a communication media. A signal is a function of data. Like data, signals are also of two types : a) Analog Signal and b) Digital Signal

Analog signals are continuous valued where as digital signals are discrete valued. Digital signal can have only a limited number of defined values i.e. 0 and 1





Signaling : it is an act of sending signal over communication media

Transmission : Communication of data by propagation and processing is known as transmission

ANALOG AND DIGITAL DATA

The term analog and digital correspond to continuous and discrete respectively. Analog data takes on continuous values on some interval.

Analog transmission is a means of transmitting analog signals regardless of their content. The dtat may be analog or digital

Digital transmission is the transfer of information through a medium in digital form. A digital signal can be transmitted only for a limited distance

Data communication is the transfer of information that is in digital form, before it enters the communication, because the original data is digital in nature

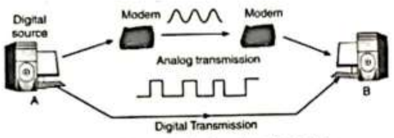


Fig. 2.3. Two methods of sending digital data

Two ways of transmitting analog information, in either cases it is not data communications, because the original information is not digital.

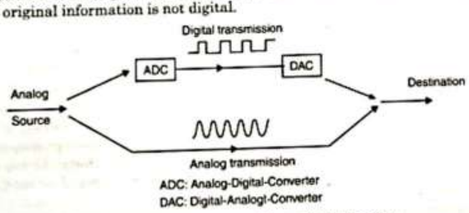


Fig. 2.4. Two ways of transmitting analog information

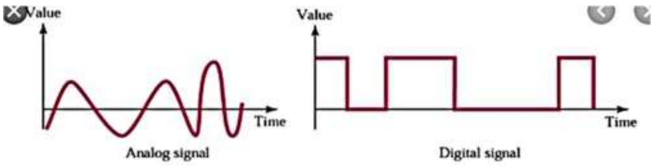
ANALOG AND DIGITAL SIGNALS

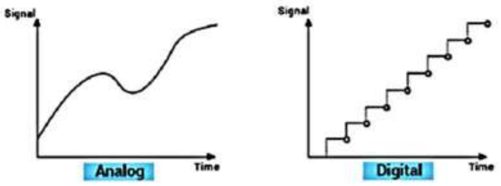
Signal is generated by a transmitter and transmitted over a medium. Signal is a function of time or frequency. A signal is any function that carries information. Based on the range of variation of independent variable, signals can be devided as : a) continuous-time (analog) and discrete-time (digital).

A signal is a function of time, but can also be expressed as a function of frequency; that is, the signal consists of components of different frequencies.

Any signal can be classified into one of the two types: Analog or Digital. An anolog signal is continuously varying signal. For instance, if we measure the room temperature continuously and plot a graph with the time on x-axis and temperature on y-axis, we get a continuous waveform.

Analog signals can have infine value in a range, where as digital signals can have only a limited value i.e. 0 and 1





The main characteristics of Analog Signals are :

- Amplitude
- Frequency
- Phase

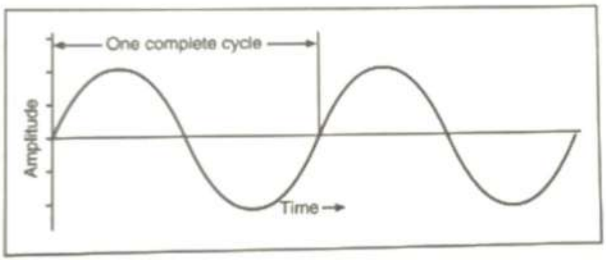
Amplitude :

This is the strength of the signal. It can be expressed a number of different ways. The higher the amplitude, the stronger the signal. The decibel is a popular measure of signal strength.

Sound level	Types of sound
40db	normal speech
90db	lawn mowers
110db	shotgun blast
120db	jet engine taking off

Frequency :

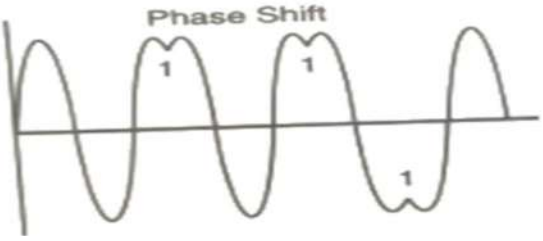
This is the rate of change the signal undergoes every second, expressed in Hz or cycles per second. A 30Hz signal changes thirty times a second. In speech, it is referred to as number of vibrations per second.



A cycle is one complete movement of the wave from its original start position and back to the same point again. The number of cycles (wages) within a one second time interval is called cycle-per-second or Hz

Phase :

This is the rate at which the signal changes its relationship to time, expressed as degree. One complete cycle of wave begins at a certain point and continues till the same point is reached again. Phase shift occurs when the cycle does not complete, and a new cycle begins before the previous one has fully completed. Phase shift is caused by imperfections in cable media, such as joins and imperfect terminations.



Periodic and Non-Periodic Signals

Both analog and digital signals can take one of two forms : periodic or non-periodic. A periodic signal completes a pattern within a measurable time frame, called “period” and repeats that pattern over subsequent identical periods. The completion of one full pattern is called CYCLE. A non-periodic signal changes without exhibiting a term or cycle that repeats over time.

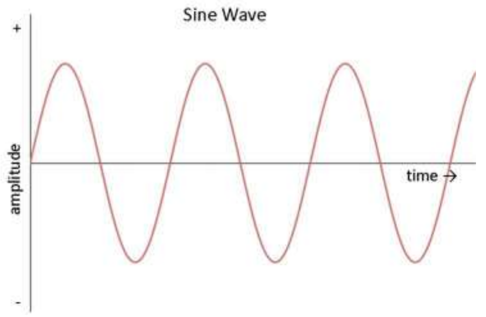
In data communication, we commonly use periodic analog signals and non-periodic digital signals.

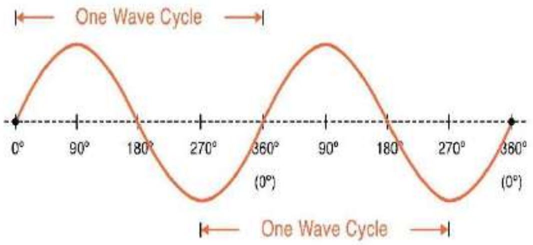
PERIODIC ANALOG SIGNALS

Periodic analog signals can be classified as simple or composite. A simple periodic analog signal, **a sine wave,** can not be decomposed into simpler signals. A composite periodic analog signal is composed of multiple sine waves

Sine Wave :

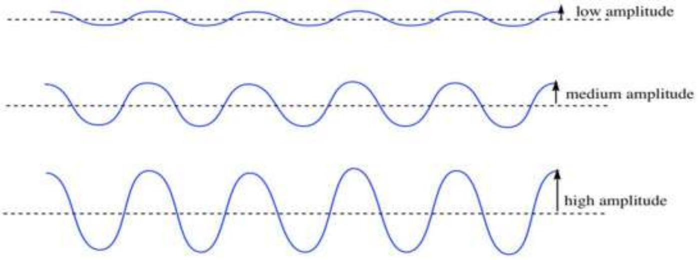
The sine wave is the most fundamental form of a periodic analog signal. Each cycle consists of a signal arc above the time axis followed by a single arc below it.





A sign wave can be represented by three parameters i.e. a) peak amplitude, b) frequency, c) phase

a) Peak Amplitude : The peak amplitude of a signal is the absolute value of its highest intensity, proportional to the energy it carries. For electric signals, peak amplitude is normally measured in volts.

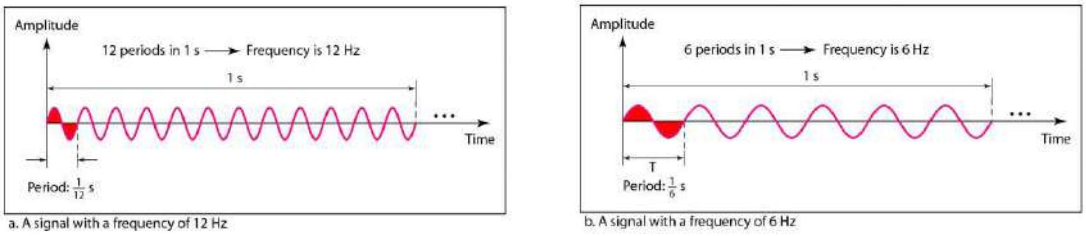


b) Period and Frequency : Period refers to the amount of time, in second, a signal needs to complete one cycle. Frequency refers to the number of periods in 1 sec. Period and Frequency are just one characteristics defined in two ways. Period is the inverse of frequency and frequency is the inverse of period. So frequency and period are inverse with each other.

$$\text{Frequency} = \frac{1}{\text{Periodic time}} \quad \text{or} \quad f = \frac{1}{\text{T}}\ \text{Hz}$$

$$\text{Periodic time} = \frac{1}{\text{Frequency}} \quad \text{or} \quad \text{T} = \frac{1}{f}\ \text{sec}$$

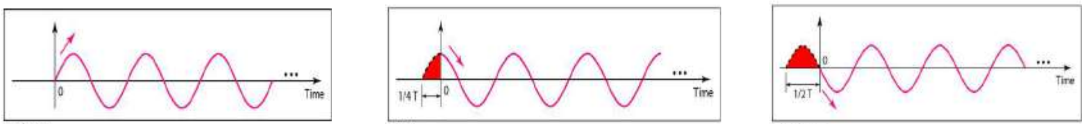
Period is formally expressed in seconds. Frequency is formally expressed in Hz, which is cycle per second. Frequency is the rate of change with respect to time. Change in short of span of time means high frequency. Change over a long span of time means low frequency.



a. A signal with a frequency of 12 Hz

b. A signal with a frequency of 6 Hz

c) Phase : Phase describes the position of the waveform relative to time 0. Phase is measured in degree or radians [360^0 is 2 π rad; 1^0 is 2π/360, and 1 rad is 360/(2π)]. A phase shift of 360^0 corresponds to a shift of a complete period; a phase shift of 180^0 corresponds to a shift of ½ of a period; and a phase shift of 90^0 corresponds to a shift of ¼ of a period.



a. 0 degrees

b. 90 degrees

c. 180 degrees

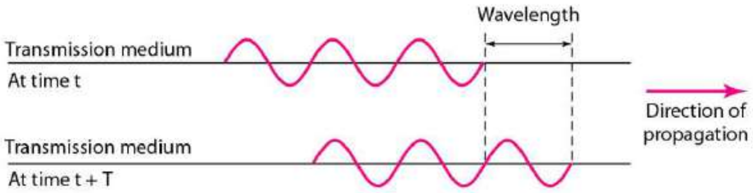
Example : A sine wave is offset 1/6 cycle with respect to time 0. What is its phase in degrees and radians?

Solution : We know that 1 complete cycle is 360°. Therefore, 1/6 cycle is

$$\frac{1}{6} \times 360 = 60° = 60 \times \frac{2\pi}{360} \text{ rad} = \frac{\pi}{3} \text{ rad} = 1.046 \text{ rad}$$

Wavelength :

Wavelength is another characteristic of a signal travelling through a transmission medium. Wavelength binds the period or the frequency of simple sine wave to the propagation speed.



While the frequency of a signal is independent of the medium, the wavelength depends on both the frequency and medium. In data communication, wavelength can be used to describe the transmission of light in an optical fiber. Wavelength is the distance, a simple signal can travel in one period.

Wavelength can be calculated if one is given the propagation speed and the period of the signal.

$$\textbf{Wavelength = propagation speed x period} = \frac{\textbf{propagation speed}}{\textbf{frequency}}$$

In a coaxial or fiber-optics cable, the wavelength is shorter (0.5micrometre - µm) due to the propagation speed in cable is decreased.

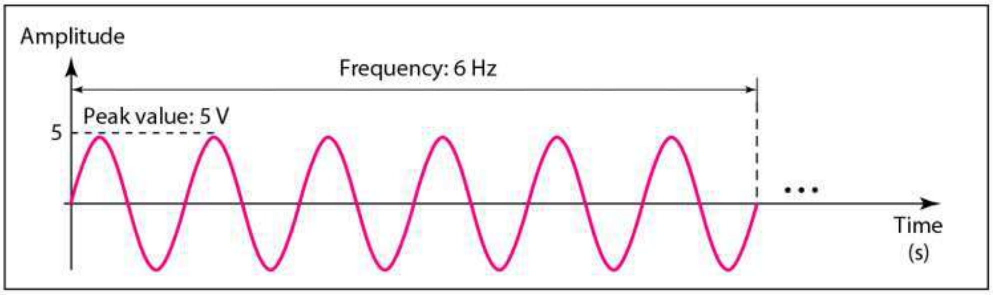
SIGNAL CHARATERISTICS :

A signal can be represented as a function of time i.e. it varies with time. It can also be expressed as a function of frequency i.e. a signal can be considered as composition of different frequency components. Therefore, a signal has both a) time-domain and b) frequency-domain

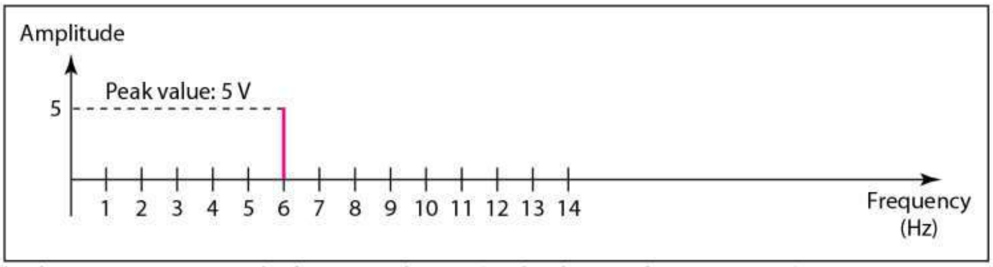
Time and Frequency Domain :

A) Time-Domain : The time-domain plot shows changes in signal amplitude with respect to time. Phase is not explicitly shown on a time-domain plot.

B) Frequency-Domain : A frequency-domain plot is concerned with only the peak value and the frequency. Changes of amplitude during one period are not shown.

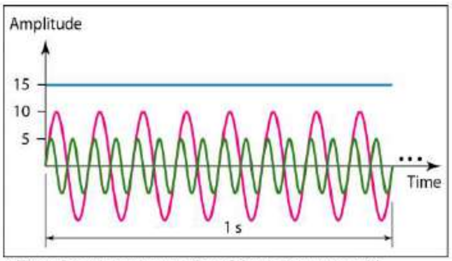


a. A sine wave in the time domain (peak value: 5 V, frequency: 6 Hz)

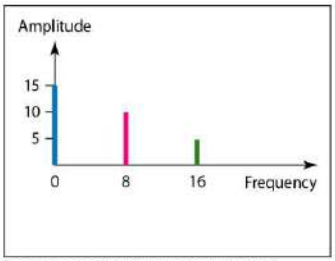


b. The same sine wave in the frequency domain (peak value: 5 V, frequency: 6 Hz)

A complete sine wave in the time domain can be represented by one single spike in the frequency domain. The frequency domain is more compact and useful when we are dealing with more than one sine wave.



a. Time-domain representation of three sine waves with frequencies 0, 8, and 16



b. Frequency-domain representation of the same three signals

The above Figure shows three sine waves, each with different amplitude and frequency. All can be represented by three spikes in the frequency domain.

A single-frequency sine wave is not useful in data communications; we need to send a composite signal, a signal made of many simple sine waves. A composite signal can be periodic or non-periodic. A periodic composite signal can be decomposed into a series of simple sine waves with discrete frequencies – frequencies that have integer values (1, 2, 3, …..). A non-periodic composite signal can be decomposed into a combination of an infinite number of simple sine waves with continuous frequencies, that have real values.

Figure 3.9 *A composite periodic signal*

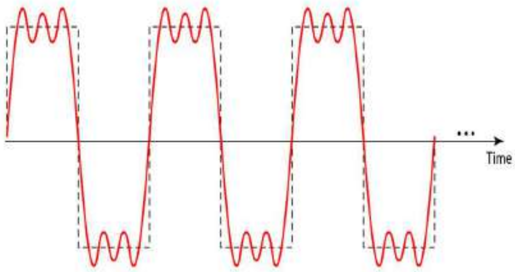
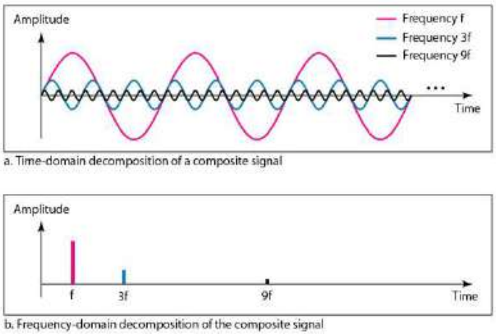


Figure 3.10 *Decomposition of a composite periodic signal in the time and frequency domains*



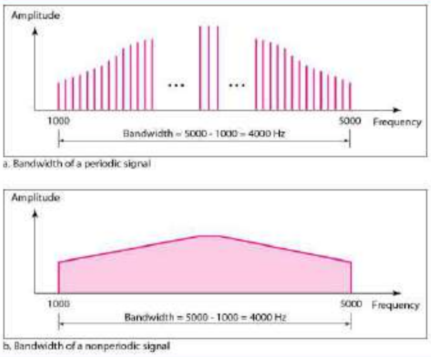
a. Time-domain decomposition of a composite signal

b. Frequency-domain decomposition of the composite signal

Bandwidth :

The range of frequencies contained in a composite signal is its bandwidth. The bandwidth of a composite signal is the difference between the highest and the lowest frequencies contained in that signal.

Example : if a composite signal contains frequencies between 1000 and 5000, its bandwidth is 4000 (5000-1000)



a. Bandwidth of a periodic signal

b. Bandwidth of a nonperiodic signal

DIGITAL SIGNALS

In addition to being represented by an analog signal, information can also be represented by a digital signal. For example, a 1 can be encoded as a positive voltage and a 0 as zero voltage. A digital signal can have more than two levels. In this case, we can send more than 1 bit for each level.

Bit Rate :

Most digital signals are non-periodic and therefore period and frequency are not appropriate characteristics. BIT Rate is used to describe digital signals. The bit rate is the number of bits sent in 1s, expressed in bits per second (bps).

Example1 :

Assume we need to download text documents at the rate of 100 pages per minute. What is the required bit rate of the channel?

Solution :

A page is an average of 24 lines with 80 characters in each line. If we assume that one character requires 8 bits, the bit rate is :

24 x 80 x 8 x 100 = 1636,000 bps = 1.636mbps

Example2 :

What is the bit rate for high-definition TV (HDTV)?

Solution :

HDTV uses digital signals to broadcast high quality video signals. The HDTV screen is normally a ratio of 16 : 9. There are 1920 by 1080 pixels per screen, and the screen is renewed 30 times per second. 24 bits represents one color pixel.

$$\boxed{1920 \times 1080 \times 30 \times 24 = 1{,}492{,}992{,}000 \text{ or } 1.5 \text{ Gbps}}$$

The TV stations reduce this rate to 20 to 40 Mbps through compression.

Bit Length :

Line wavelength for analog signal (the distance one cycle occupies on the transmission medium), we can use **bit length** for digital signals. The bit length is the distance one bit occupies on the transmission medium.

Bit length = propagation speed x bit duration

Digital Signal as a Composite Analog Signal

What is a Fourier Analysis? (tutorialspoint.com)

Based on Fourier analysis, a digital signal is a composite analog signal. The bandwidth is infinite. In the time domain, a digital signal comprises connected vertical and horizontal line segments. A vertical time domain means a frequency of infinity (sudden change in time); a horizontal line in the time domain means a frequency of 0 (no change in time).

If the digital signal is periodic, which is rare in data communication, the decomposed signal has a frequency domain representation with an infinite bandwidth and discrete frequencies. If the digital signal is non-periodic, the decomposed signal still has an infinite bandwidth, but the frequencies are continuous.

Transmission of Digital Signals :

Digital transmission is different from analog transmission. Digital signal is a series of discrete pulses, which represents either 1 bit or 0 bit. Every computer uses a coding scheme that defines what combinations of 1s and 0s constitute all the characters in a character set.

In electrical networks, 1 bits are represented as high voltage and 0 bits represented as absence / null /low voltage. In optical networks, 1 bits are represented by presence of light and 0 bits are represented as absence of light. Then 1s and 0s / On/Off are carried out through the network. The receiving device repackages the 1s and 0s to determine what character is being represented.

As a digital pulse travels over a distance, it may losses power, similar to an analog signal, such weakened and impaired signal enters the regenerative repeater, where the repeater examines the signal to determine what was supposed to be a 1 and 0. The repeater regenerates a new signal to pass on to the next point in the network to eliminate noise.

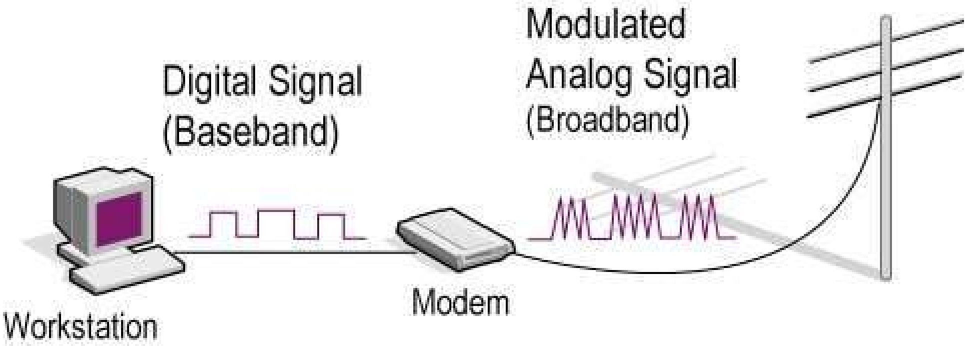
Baseband Transmission :

Baseband Transmission is a signaling technology that sends digital signals over a single frequency as discrete electrical pulses. The entire bandwidth of a baseband system carries only one data signal and is generally less than the amount of bandwidth available on a broadband transmission system.

The baseband signal is bidirectional so that a baseband system can both transmit and receive signals simultaneously. Baseband signals can be regenerated using repeaters in order to travel longer distances before weakening and becoming unusable because of attenuation.

Baseband transmission technologies do not use modulation, but they can use time-division multiplexing (TDM) to accommodate multiple channels over a single baseband transmission line.

Common local area network (LAN) networking technologies such as Ethernet use baseband transmission technology. All stations on a baseband network share the same transmission medium, and they use the entire bandwidth of that medium for transmission. As a result, only one device on a baseband network can transmit at a given instant, resulting in the need for a media access control method to handle contention.



Baseband transmission means sending a digital signal over a channel without changing the digital signal to an analog signal. When low frequencies are sent, they would not reach the final destination. These low frequency signals are called baseband signals and once they are change to higher frequencies, the name changes to RF (radio frequency) signals.

Baseband transmission –

1. Digital signalling.
2. Frequency division multiplexing is not pssible.
3. Baseband is bi-directional transmission.

4. Short distance signal travelling.
5. Entire bandwidth is for single signal transmission.
6. Example: Ethernet is using Basebands for LAN.

Bradband Transmission :

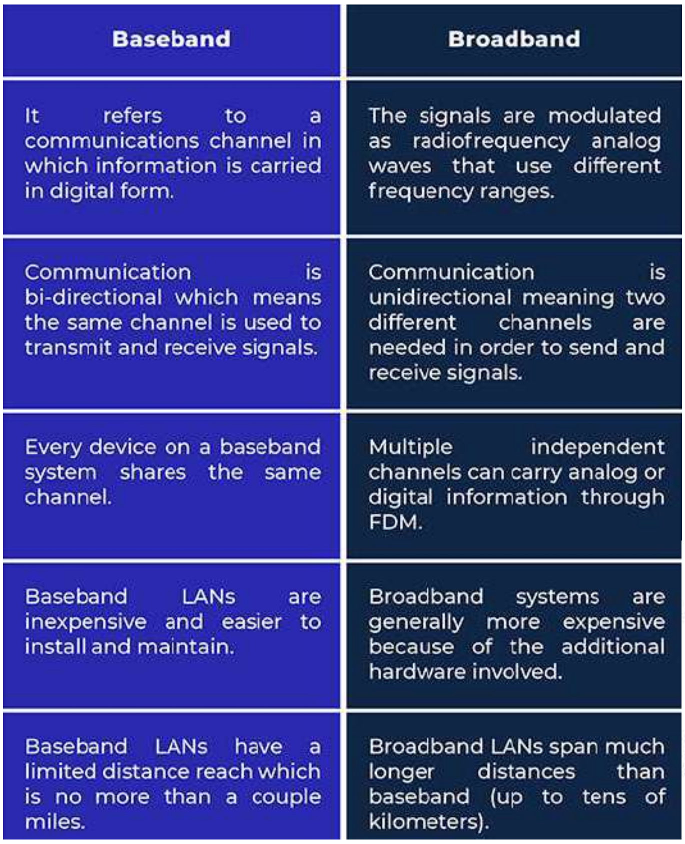
Broadband system use modulation techniques to reduce the effect of noise in the environment. Broadband transmission employs multiple channel unidirectional transmission using combination of phase and amplitude modulation.

Broadband Transmission is a signaling technology that sends signals simultaneously over a range of different frequencies as electromagnetic waves. The bandwidth of a broadband system can usually carry multiple, simultaneous data signals.

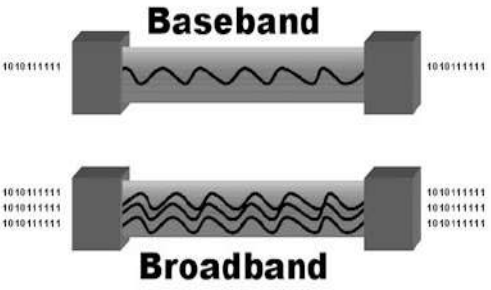
These signals are unidirectional – traveling in only one direction at a time – so a broadband system can generally either transmit or receive but cannot do both simultaneously. Broadband signals can be regenerated using amplifiers in order to travel longer distances before becoming attenuated.

Broadband transmission –

1. Analog signalling.
2. Transmission of data is unidirectional.
3. Signal travelling distance is long.
4. Frequency division multiplexing possible.
5. Simultaneous transmission of multiple signals over different frequencies.
6. Example : Used to transmit cable TV to premises.



Baseband	Broadband
It refers to a communications channel in which information is carried in digital form.	The signals are modulated as radiofrequency analog waves that use different frequency ranges.
Communication is bi-directional which means the same channel is used to transmit and receive signals.	Communication is unidirectional meaning two different channels are needed in order to send and receive signals.
Every device on a baseband system shares the same channel.	Multiple independent channels can carry analog or digital information through FDM.
Baseband LANs are inexpensive and easier to install and maintain.	Broadband systems are generally more expensive because of the additional hardware involved.
Baseband LANs have a limited distance reach which is no more than a couple miles.	Broadband LANs span much longer distances than baseband (up to tens of kilometers).



How does Broadband Transmission work?

Broadband transmissions are divided into multiple bands or channels by multiplexers using a multiplexing scheme such as frequency-division multiplexing (FDM). Each channel has a carrier frequency that is modulated to carry the signal from a given source. At the receiving station, multiplexers separate the various signals. Guard bands are used to prevent interference among channels.

Broadband transmission is typically used for environments in which video, audio, and data need to be transmitted simultaneously. Cable television systems are based on broadband transmission technologies. Other examples of broadband services include T-carrier services, Asynchronous Transfer Mode (ATM), and variants of Digital Subscriber Line (DSL).

TRANSMISSION IMPAIRMENT

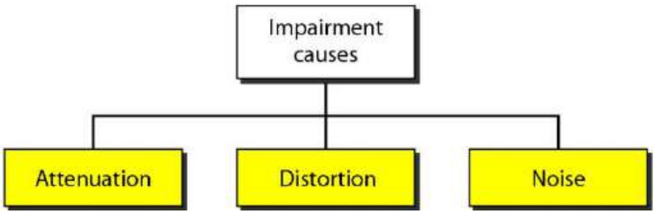
Transmission impairment is a property of a transmission medium which causes the signal to be degraded, reduced in amplitude, distorted. Impairment can introduce errors into digital signals. Signals travel through transmission media, which are not perfect. The imperfection causes signal impairment. This means, the signal at the beginning of the medium is not the same at the end of the medium.

The source of transmission impairment are :

a) Maximum length of network link
b) Choice of physical transmission media
c) Choice of encoding method
d) Data rate supported by the medium

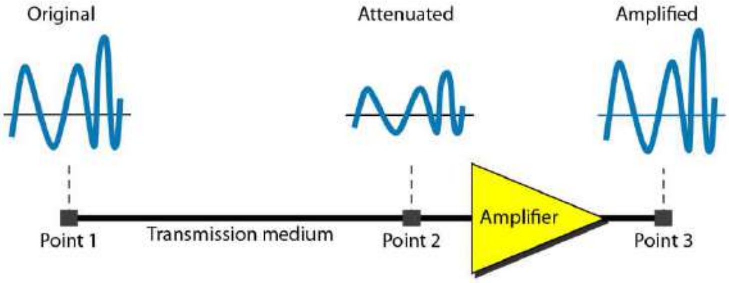
Transmission lines suffers three major problems :

a) Attenuation
b) Delay distortion
c) Noise.



a) Attenuation :

Attenuation means a loss of energy. When a signal (simple or composite) travels through a medium, it loses some of its energy in overcoming the resistances of the medium. That is why a wire carrying electric signals gets warm. Some of the electrical energy in the signal converted into heat. To compensate this loss, amplifiers are used to amplify the signal.

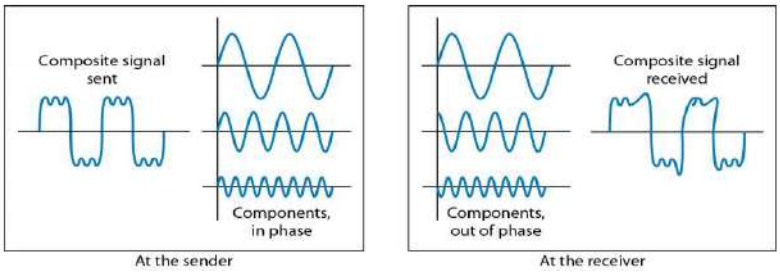


To know that a signal has lost/gained strength, one unit is used i.e. **decibel.** The decibel (dB) measures the relative strengths of two / one signal at two different points. The dB is negative (-ve) if a signal is attenuated and positive (+ve) if a signal s amplified

dB = 10log$_{10}\frac{P2}{P1}$, where, P1 and P2 are the powers of a signal at points 1 and 2

b) Delay Distortion :

Distortion means that the signal changes its form/shape. It can occur in a composite signal made of different frequencies. Each signal component has its own propagation speed through a medium and therefore, its own delay in arriving at the destination. Differences in delay may create a difference in phase if the delay s not exactly the same as the period duration. In other words, signal components at the receiver have phase different from what they had at the sender. The shape of the composite signal is therefore, not the same.



c) Noise :

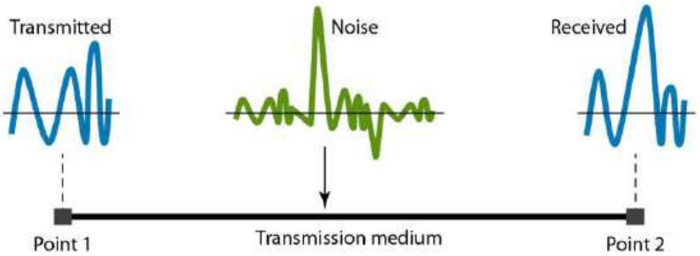
Noise s another cause of impairment. Several types of noise may corrupt the signal like : thermal noise, induced noise, crosstalk, impulse noise.

Thermal Noise – it is the random motion of electrons in a wire which creates an extra signal, which is not originally sent by the sender.

Induced Noise – it comes from sources such as : motors and appliances. These devices act as sending antenna and the transmission medium acts as the receiving antenna.

Crosstalk – it is the effect of one wire on the other. One wire acts as a sending antenna and other wire acts as the receiving antenna.

Impulse Noise – it is a spike (a signal with high energy in a very short time) that comes from power lines, lightning, etc.



Signal-to-Noise Ratio (SNR) :

To find the theoretical bit rate limit, we need to know the ratio of the signal power to the noise power. The SNR is defined as

$$\textbf{SNR} = \frac{\textit{average signal power}}{\textit{average noise power}}$$

CHANNEL CAPACITY

Maximum Data Rate (channel capacity) for Noiseless and Noisy channels. Data rate governs the speed of data transmission. A very important consideration in data communication is how fast we can send data, in bits per second, over a channel. Data rate depends upon 3 factors:

a) bandwidth available
b) Number of levels in digital signal
c) The quality of the channel – level of noise

Two theoretical formulas were developed to calculate the data rate:

1) by Nyquist for a noiseless channel 2) by Shannon for a noisy channel.

1. **Noiseless Channel : Nyquist Bit Rate –**

For a noiseless channel, the Nyquist bit rate formula defines the theoretical maximum bit rate

BitRate = 2 * Bandwidth * log2(L)

In the above equation, bandwidth is the bandwidth of the channel, L is the number of signal levels used to represent data, and BitRate is the bit rate in bits per second.

Bandwidth is a fixed quantity, so it cannot be changed. Hence, the data rate is directly proportional to the number of signal levels.

Note –Increasing the levels of a signal may reduce the reliability of the system.

Input1 : Consider a noiseless channel with a bandwidth of 3000 Hz transmitting a signal with two signal levels. What can be the maximum bit rate?

Output1 : BitRate = 2 * 3000 * log2(2) = 6000bps

Input2 : We need to send 265 kbps over a noiseless channel with a bandwidth of 20 kHz. How many signal levels do we need?

Output2 : 265000 = 2 * 20000 * log2(L)

log2(L) = 6.625

L = 26.625 = 98.7 levels

2. **Noisy Channel : Shannon Capacity –**

In reality, we cannot have a noiseless channel; the channel is always noisy. Shannon capacity is

used, to determine the theoretical highest data rate for a noisy channel:

Capacity = bandwidth * log2(1 + SNR)

In the above equation, bandwidth is the bandwidth of the channel, SNR is the signal-to-noise ratio, and capacity is the capacity of the channel in bits per second.

Bandwidth is a fixed quantity, so it cannot be changed. Hence, the channel capacity is directly proportional to the power of the signal, as SNR = (Power of signal) / (power of noise).

The signal-to-noise ratio (S/N) is usually expressed in decibels (dB) given by the formula:

10 * log10(S/N)

so for example a signal-to-noise ratio of 1000 is commonly expressed as:

10 * log10(1000) = 30 dB.

Examples:

Input1 : A telephone line normally has a bandwidth of 3000 Hz (300 to 3300 Hz) assigned for data communication. The SNR is usually 3162. What will be the capacity for this channel?

Output1 : C = 3000 * log2(1 + SNR) = 3000 * 11.62 = 34860 bps

Input2 : The SNR is often given in decibels. Assume that SNR(dB) is 36 and the channel bandwidth is 2 MHz. Calculate the theoretical channel capacity.

Output2 : SNR(dB) = 10 * log10(SNR)

SNR = 10(SNR(dB)/10)

SNR = 103.6 = 3981

Hence, C = 2 * 106 * log2(3982) = 24 MHz

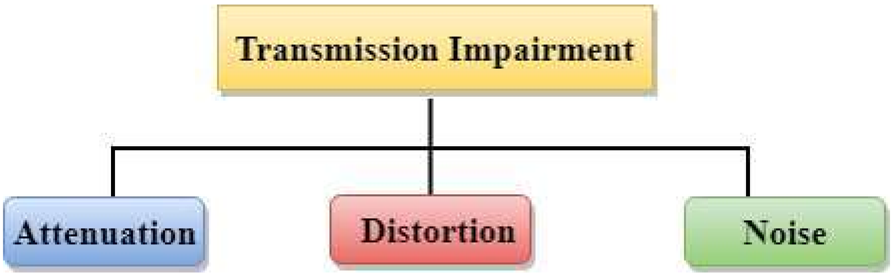
TRANSMISSION MEDIA

- Transmission media is a communication channel that carries the information from the sender to the receiver. Data is transmitted through the electromagnetic signals.
- The main functionality of the transmission media is to carry the information in the form of bits through **LAN**(Local Area Network).
- It is a physical path between transmitter and receiver in data communication.
- In a copper-based network, the bits in the form of electrical signals.
- In a fibre based network, the bits in the form of light pulses.
- In **OSI**(Open System Interconnection) phase, transmission media supports the Layer 1. Therefore, it is considered to be as a Layer 1 component.
- The electrical signals can be sent through the copper wire, fibre optics, atmosphere, water, and vacuum.
- The characteristics and quality of data transmission are determined by the characteristics of medium and signal.
- Transmission media is of two types are wired media and wireless media. In wired media, medium characteristics are more important whereas, in wireless media, signal characteristics are more important.
- Different transmission media have different properties such as bandwidth, delay, cost and ease of installation and maintenance.
- The transmission media is available in the lowest layer of the OSI reference model, i.e., **Physical layer**.

Some factors need to be considered for designing the transmission media:

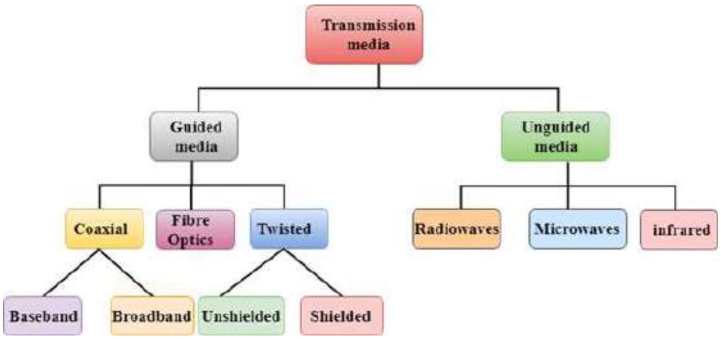
- **Bandwidth:** All the factors are remaining constant, the greater the bandwidth of a medium, the higher the data transmission rate of a signal.
- **Transmission impairment:** When the received signal is not identical to the transmitted one due to the transmission impairment. The quality of the signals will get destroyed due to transmission impairment.
- **Interference:** An interference is defined as the process of disrupting a signal when it travels over a communication medium on the addition of some unwanted signal.

Causes Of Transmission Impairment:



- **Attenuation:** Attenuation means the loss of energy, i.e., the strength of the signal decreases with increasing the distance which causes the loss of energy.
- **Distortion:** Distortion occurs when there is a change in the shape of the signal. This type of distortion is examined from different signals having different frequencies. Each frequency component has its own propagation speed, so they reach at a different time which leads to the delay distortion.
- **Noise:** When data is travelled over a transmission medium, some unwanted signal is added to it which creates the noise.

CLASSIFICATION OF TRANSMISSION MEDIA:



GUIDED MEDIA

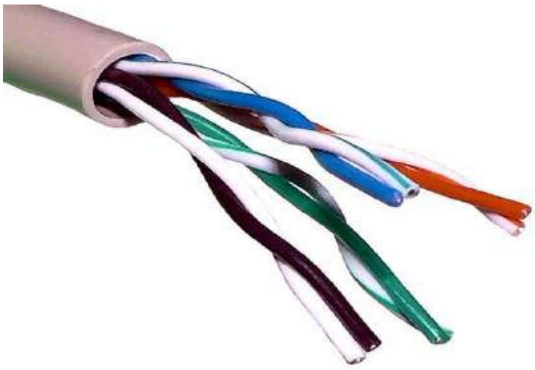
It is defined as the physical medium through which the signals are transmitted. It is also known as Bounded media.

Types Of Guided media:

Twisted pair:

Twisted pair is a physical media made up of a pair of cables twisted with each other. 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.

A twisted pair cable is economical as compared to other transmission media. Installation of the twisted pair cable is easy, and it is a lightweight cable. The frequency range for twisted pair cable is from 0 to 3.5KHz.



There are two types of twisted pair cables:

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

STP :

A shielded twisted pair is a cable that contains the mesh surrounding the wire that allows the higher transmission rate.

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

Characteristics Of Shielded Twisted Pair:

- The cost of the shielded twisted pair cable is not very high and not very low.
- An installation of STP is easy.
- It has higher capacity as compared to unshielded twisted pair cable.
- It has a higher attenuation.
- It is shielded that provides the higher data transmission rate.

Application of Twisted Pair:

- In telephone lines to carry voice and data channels
- In local loop
- In DSL
- LANs like:10 base T, 100 base T
- In ISDN

Disadvantages

- It is more expensive as compared to UTP and coaxial cable.
- It has a higher attenuation rate.

UTP: An unshielded twisted pair is widely used in telecommunication. 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.

Following are the categories of the unshielded twisted pair cable:

- **Category 1:** Category 1 is used for telephone lines that have low-speed data.
- **Category 2:** It can support upto 4Mbps.
- **Category 3:** It can support upto 16Mbps.
- **Category 4:** It can support upto 20Mbps. Therefore, it can be used for long-distance communication.
- **Category 5:** It can support upto 200Mbps.

Advantages Of Unshielded Twisted Pair:

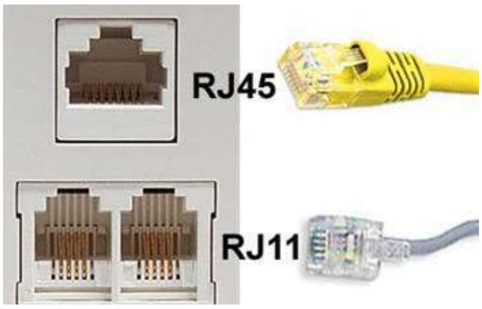
- It is cheap.
- Installation of the unshielded twisted pair is easy.
- It can be used for high-speed LAN.

Disadvantage:

- This cable can only be used for shorter distances because of attenuation.

Connectors for Twisted Pair Cable :

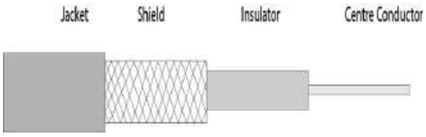
The UTP cable is the most commonly used cable in computer network. The most common UTP connector is rJ-45. It is a male-female type keyed connector which can be inserted in only one way.

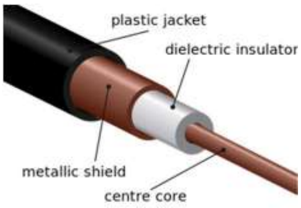


Coaxial Cable

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.

- Coaxial cable is very commonly used transmission media, for example, TV wire is usually a coaxial cable.
- The name of the cable is coaxial as it contains two conductors parallel to each other.
- It has a higher frequency as compared to Twisted pair cable.
- The inner conductor of the coaxial cable is made up of copper, and the outer conductor is made up of copper mesh. The middle core is made up of non-conductive cover that separates the inner conductor from the outer conductor.
- The middle core is responsible for the data transferring whereas the copper mesh prevents from the **EMI**(Electromagnetic interference).





Because of its structure,the coaxial 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.

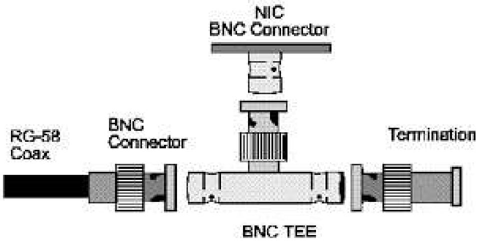
Coaxial cable is of two types:

1. **Baseband transmission:** It is defined as the process of transmitting a single signal at high speed.
2. **Broadband transmission:** It is defined as the process of transmitting multiple signals simultaneously.

Connectors Of Coaxial Cable :

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

BNC connector: The BNC connector is a miniature quick connect/disconnect radio frequency connector used for coaxial cable.



BNC T-Connector : A tee connector is an electrical connector that connects three cables together. It is usually in the shape of a capital T.

Tee connectors can be used to split radio frequency power from a cable into two. They can be used to attach a piece of electronic test equipment. Tee connectors were much used on 10M Ethernet over coax networks.

BNC Terminator : BNC terminator absorbs the electrical energy of the signal as it reaches the ends of cable and avoids reflection of signals. So it doesn't become noise.

If termination is missing, or if there is a break in the cable, the AC signal on the bus is reflected, rather than dissipated/dissolved, when it reaches the end. This reflected signal is indistinguishable from a collision, so no communication can take place.

Advantages Of Coaxial cable:

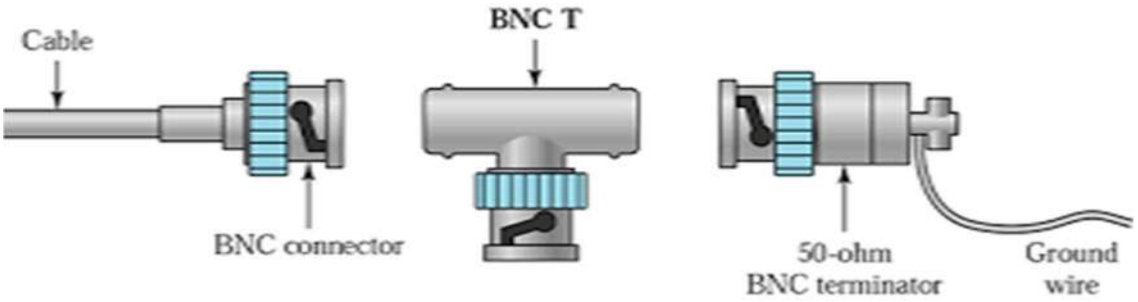
- The data can be transmitted at high speed.
- It has better shielding as compared to twisted pair cable.
- It provides higher bandwidth.

Disadvantages Of Coaxial cable:

- It is more expensive as compared to twisted pair cable.
- If any fault occurs in the cable causes the failure in the entire network.

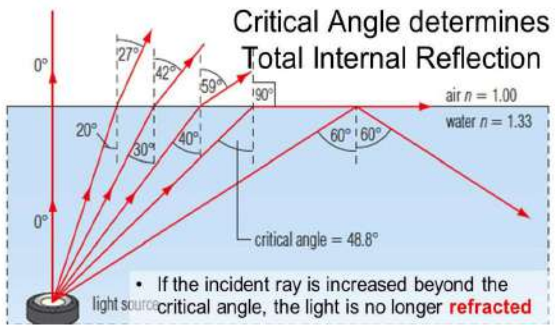
Application of Coaxial cable:

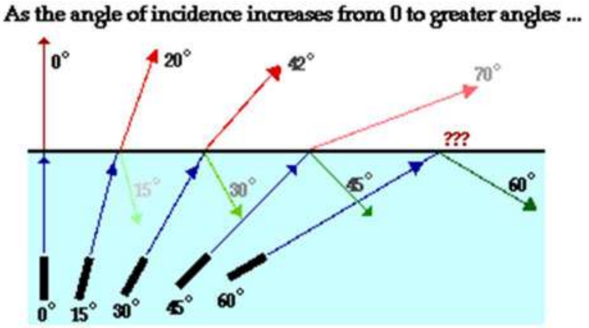
- Analog telephone network
- Digital telephone network
- Cable TV
- Traditional Ethernet LAN
- Digital transmission
- Thick and Thin Eathernet

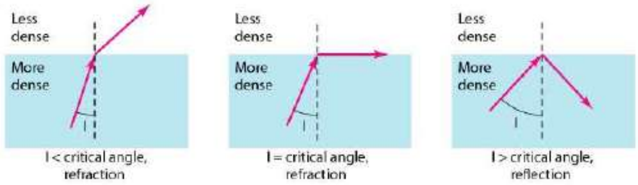


Fibre Optic

A fibre optic cable is made of glass or plastic and transmits signals in the form of light. Fibre Optic works on the properties of light. Light travels in a straight line as long as it is moving through a single uniform substance. If light rays travelling in substance suddenly enters into another substance (different density), the ray changes its direction.

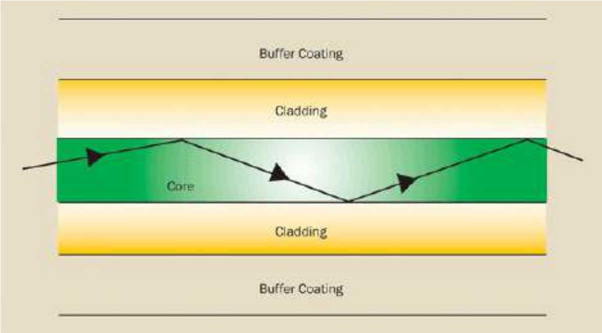

Critical Angle determines
Total Internal Reflection
0°
27°
42°
59°
90°
air n = 1.00
water n = 1.33
20°
30°
40°
60°
60°
0°
critical angle = 48.8°
• If the incident ray is increased beyond the critical angle, the light is no longer refracted
light source


As the angle of incidence increases from 0 to greater angles ...
0°
20°
42°
70°
???
15°
30°
45°
60°
0°
15°
30°
45°
60°


Less dense
More dense
I < critical angle, refraction
Less dense
More dense
I = critical angle, refraction
Less dense
More dense
I > critical angle, reflection

If the angle of incidence "I" is less than the critical angle, the ray refracted and moves closer to the surface. If the angle of incidence is equal to the critical angle, the light bends along the interface. If the angle is greater than the critical angle, the ray reflects (return) and travels again in the denser substance. The critical angle is a property of the substance and its value differs from one substance to another.

Optical fibre use reflection to guide light through a channel. A glass or plastic core is surrounded by cladding of less dense glass or plastic. The difference is density of the two materials must be such that a beam of light moving through the core is reflected off the cladding instead of being refracted into it.



When light ray hits at critical angle it tends to refracts at 90 degrees. This property has been used in fibre optic. The core of fibre 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.

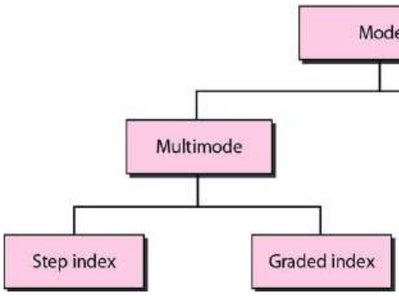
- Fibre optic cable is a cable that uses electrical signals for communication.
- Fibre optic is a cable that holds the optical fibres coated in plastic that are used to send the data by pulses of light.
- The plastic coating protects the optical fibres from heat, cold, electromagnetic interference from other types of wiring.
- Fibre optics provide faster data transmission than copper wires.

Basic elements of Fibre optic cable:

- **Core:** The optical fibre consists of a narrow strand of glass or plastic known as a core. A core is a light transmission area of the fibre. The more the area of the core, the more light will be transmitted into the fibre.
- **Cladding:** The concentric layer of glass is known as cladding. The main functionality of the cladding is to provide the lower refractive index at the core interface as to cause the reflection within the core so that the light waves are transmitted through the fibre.
- **Jacket:** The protective coating consisting of plastic is known as a jacket. The main purpose of a jacket is to preserve the fibre strength, absorb shock and extra fibre protection.

Propagation Mode :

Fibre Optic provides the highest mode of speed. It comes in two modes i.e. a) multimode mode fibre and b) single fibre. Single mode fibre can carry a single ray of light whereas multimode is capable of carrying multiple beams of light. Multimode mode can be implemented in two forms i.e. i) step-index and ii) graded-index

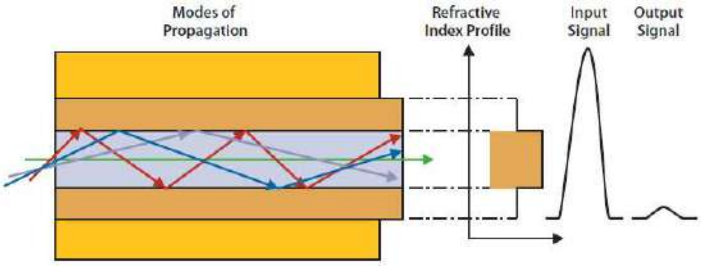


a) **Multi mode :**

In Multimode multiple beams moves through the core in different paths. In multimode, signal can be implemented in : i)multimode step index fibre and ii) multimode graded-index fibre.

i) Multimode Step-Index Fibre : In a step index fiber, the light rays propagate in zig-zag manner inside the core. The term step-index refers to the sudden change in refractive index at the boundary of core and cladding. In this method, the density of the core remains constant from the center to the edge. A light beam moves in a straight line until it reaches the interface of the core and cladding.

The number of reflections that a beam undergoes, depends on the angle of incidence of that beam. So, at the destination, all the beams does not reach simultaneously. This leads to diffusion of signal at the receiver. The step index multimode fibre are therefore not used for long distance communicaton.

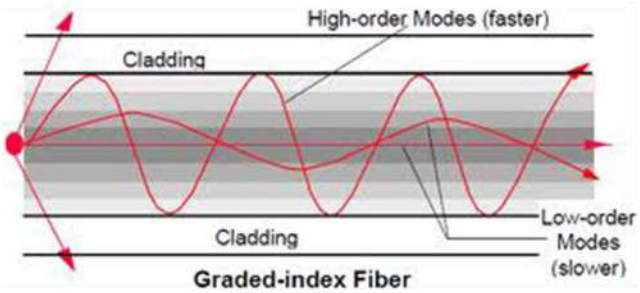


Light propagation though step-index multimode fiber

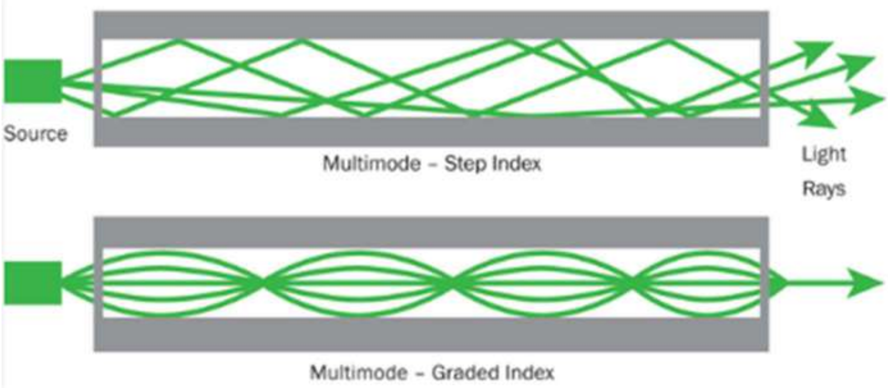
ii) Multimode Graded-Index Fibre :

It decreases the distortion / alteration of the signal through the cable. The word index refers to the index of refraction. The index of refraction is related to density. As the index of refraction is related to the density, a graded index fibre has varying density.

The refractive index of the fibre core does not remain constant throughout its bulk. It is maximum at the centre of the core and reduces gradually towards the walls of the core.



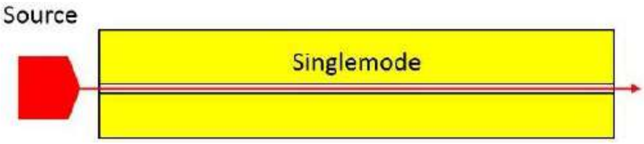
To get this type of index profile, the material in the fibre core is modified. Due to the modifications in the index profile, the light gets refracted inside the fibre card and does not travel in straight line

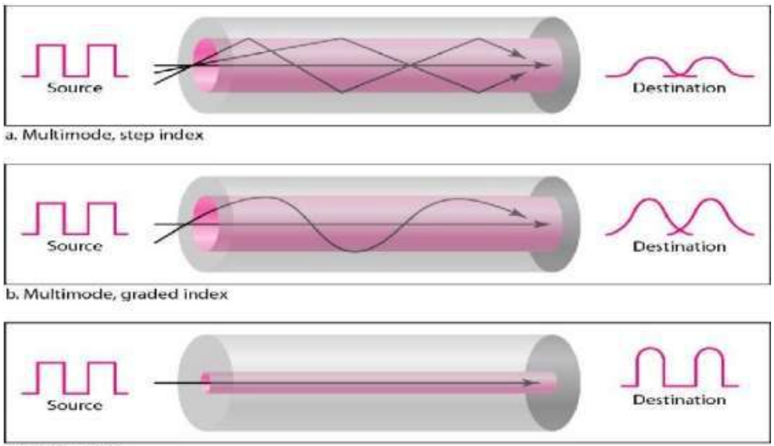


b) Single mode :

Single mode uses step index fibre and a highly focused source of light that limits beams to a small range of angles, all close to the horizontal. The single mode fibre is of smaller diameter that multimode fibre.

It has lower density i.e. index of refraction. The decrease in density results in a critical angle that is close enough to 90^0 to make the propagation of beams almost horizontal. In this case, propagation of different beams is almost identical and delays are negligible. All of the beams arrive at the destination together and can be recombined without distortion to the signal.





a. Multimode, step index

b. Multimode, graded index

c. Single mode

Following are the advantages of fibre optic cable over copper:

- **Greater Bandwidth:** The fibre optic cable provides more bandwidth as compared copper. Therefore, the fibre optic carries more data as compared to copper cable.
- **Faster speed:** Fibre optic cable carries the data in the form of light. This allows the fibre optic cable to carry the signals at a higher speed.
- **Longer distances:** The fibre optic cable carries the data at a longer distance as compared to copper cable.
- **Better reliability:** The fibre optic cable is more reliable than the copper cable as it is immune to any temperature changes while it can cause obstruct in the connectivity of copper cable.
- **Thinner and Sturdier:** Fibre optic cable is thinner and lighter in weight so it can withstand more pull pressure than copper cable.

Disadvantages of Optical Fiber:

There are some disadvantages in the use of optical fiber.

- Installation and maintenance.
- Unidirectional light propagation.
- Cost.

Applications

- Fiber-optic cable is often found in backbone networks because its wide bandwidth is cost-effective. Today, with wavelength-division multiplexing (WDM), we can transfer data at a rate of 1600 Gbps.
- Some cable TV companies use a combination of optical fiber and coaxial cable, thus creating a hybrid network. Optical fiber provides the backbone structure while coaxial cable provides the connection to the user premises. This is a cost-effective configuration since the narrow bandwidth requirement at the user end does not justify the use of optical fiber. Local-area networks such as 100Base-FX network (Fast Ethernet) and 1000Base-X also use fiber-optic cable.

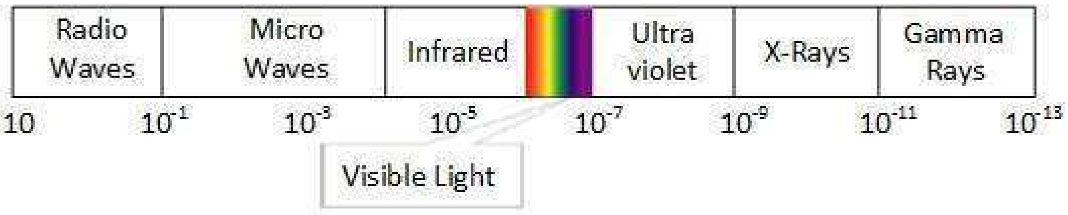
UNGUIDED MEDIA

- An unguided transmission transmits the electromagnetic waves without using any physical medium. Therefore it is also known as **wireless transmission**.
- In unguided media, air is the media through which the electromagnetic energy can flow easily.

Wireless transmission is a form of unguided media. Wireless signals are spread over in the air and are received and interpreted by appropriate antennas.

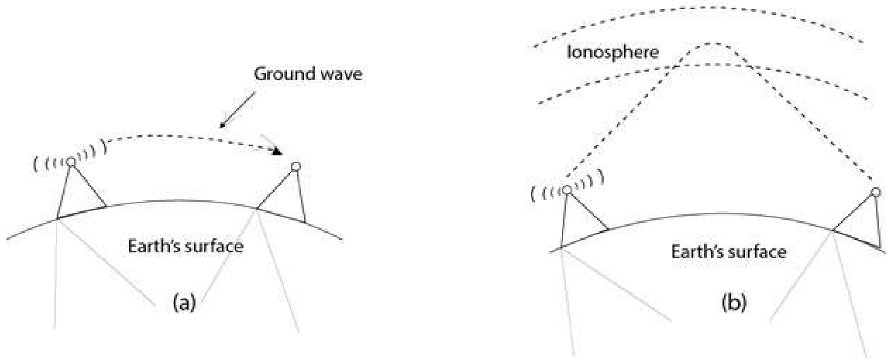
When an antenna is attached to electrical circuit of a computer or wireless device, it converts the digital data into wireless signals and spread all over within its frequency range. The receptor on the other end receives these signals and converts them back to digital data.

A little part of electromagnetic spectrum can be used for wireless transmission.



Radio Transmission

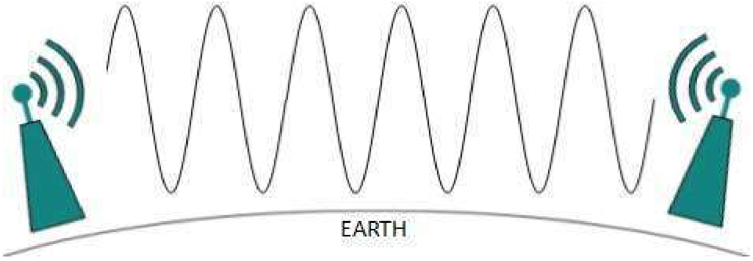
- Radio waves are the electromagnetic waves that are transmitted in all the directions of free space.
- Radio waves are omnidirectional, i.e., the signals are propagated in all the directions.
- The range in frequencies of radio waves is from 3Khz to 1 khz.
- In the case of radio waves, the sending and receiving antenna are not aligned, i.e., the wave sent by the sending antenna can be received by any receiving antenna.
- An example of the radio wave is **FM radio**.



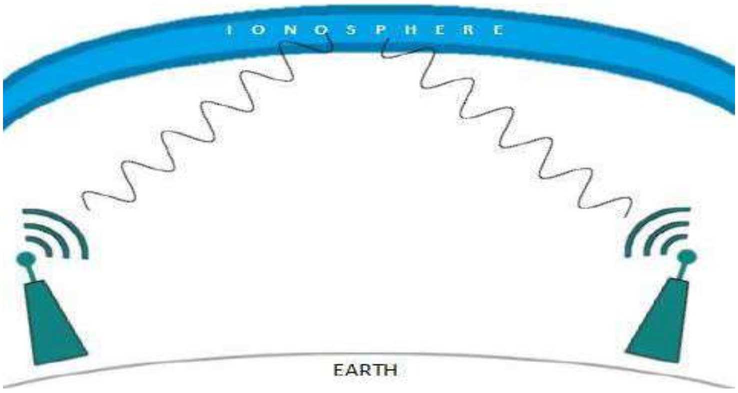
Radio frequency is easier to generate and because of its large wavelength it can penetrate through walls and structures alike. Radio waves can have wavelength from 1 mm – 100,000 km and have frequency ranging from 3 Hz (Extremely Low Frequency) to 300 GHz (Extremely High Frequency). Radio frequencies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



Applications Of Radio waves:

- A Radio wave is useful for multicasting when there is one sender and many receivers.
- An FM radio, television, cordless phones are examples of a radio wave.

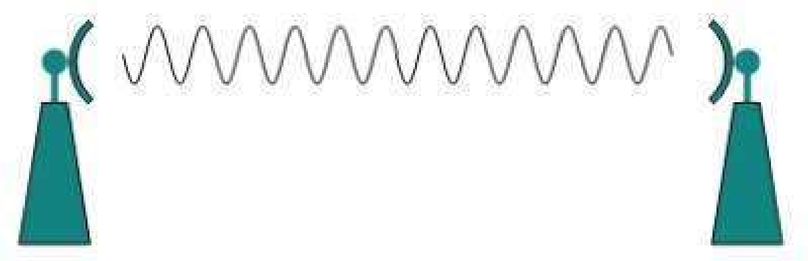
Advantages Of Radio transmission:

- Radio transmission is mainly used for wide area networks and mobile cellular phones.
- Radio waves cover a large area, and they can penetrate the walls.
- Radio transmission provides a higher transmission rate.

Microwave Transmission

Electromagnetic waves above 100 MHz 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 1 mm – 1 meter and frequency ranging from 300 MHz to 300 GHz.



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.

Microwaves are of two types:

- Terrestrial microwave
- Satellite microwave communication.

Terrestrial Microwave Transmission

- Terrestrial Microwave transmission is a technology that transmits the focused beam of a radio signal from one ground-based microwave transmission antenna to another.
- Microwaves are the electromagnetic waves having the frequency in the range from 1GHz to 1000 GHz.
- Microwaves are unidirectional as the sending and receiving antenna is to be aligned, i.e., the waves sent by the sending antenna are narrowly focussed.
- In this case, antennas are mounted on the towers to send a beam to another antenna which is km away.
- It works on the line of sight transmission, i.e., the antennas mounted on the towers are the direct sight of each other.

Characteristics of Terrestrial Microwave:

- **Frequency range:** The frequency range of terrestrial microwave is from 4-6 GHz to 21-23 GHz.
- **Bandwidth:** It supports the bandwidth from 1 to 10 Mbps.
- **Short distance:** It is inexpensive for short distance.
- **Long distance:** It is expensive as it requires a higher tower for a longer distance.
- **Attenuation:** Attenuation means loss of signal. It is affected by environmental conditions and antenna size.

Advantages Of Terrestrial Microwave:

- Microwave transmission is cheaper than using cables.
- It is free from land acquisition as it does not require any land for the installation of cables.
- Microwave transmission provides an easy communication in terrains as the installation of cable in terrain is quite a difficult task.
- Communication over oceans can be achieved by using microwave transmission.

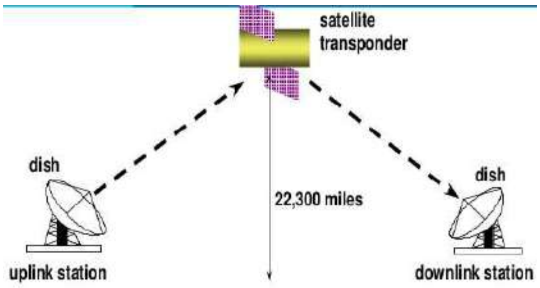
Disadvantages of Terrestrial Microwave transmission:

- **Eavesdropping:** An eavesdropping creates insecure communication. Any malicious user can catch the signal in the air by using its own antenna.
- **Out of phase signal:** A signal can be moved out of phase by using microwave transmission.
- **Susceptible to weather condition:** A microwave transmission is susceptible to weather condition. This means that any environmental change such as rain, wind can distort the signal.
- **Bandwidth limited:** Allocation of bandwidth is limited in the case of microwave transmission.

Satellite Microwave Communication

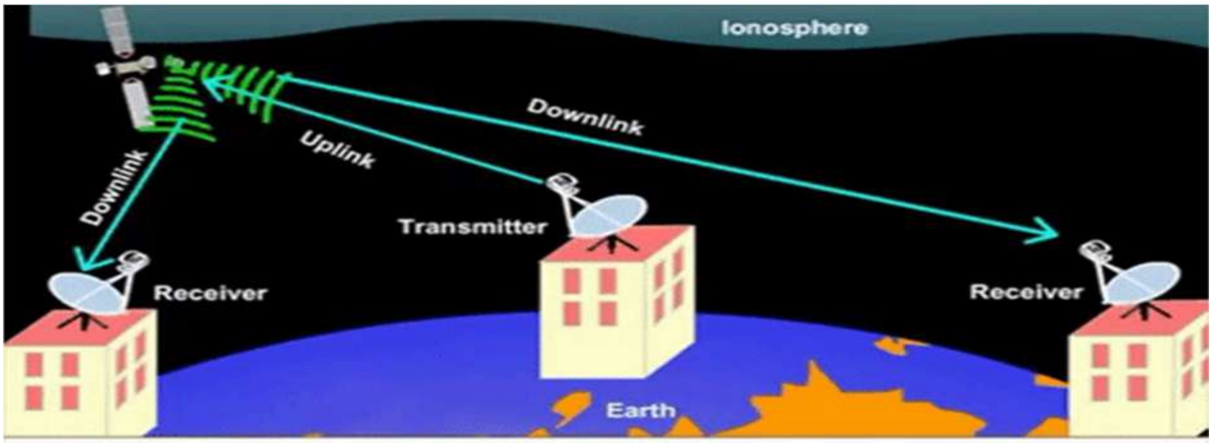
Satellite Microwave Transmission System uses satellites for broadcasting and receiving of signals. These systems need satellites which are in the geostationary orbit which is 36000 km above the earth.
The satellites operate as repeaters with receiving antenna, transponder and transmitting of signals

- A satellite is a physical object that revolves around the earth at a known height.
- Satellite communication is more reliable nowadays as it offers more flexibility than cable and fibre optic systems.
- We can communicate with any point on the globe by using satellite communication.



How Does Satellite work?

The satellite accepts the signal that is transmitted from the earth station, and it amplifies the signal. The amplified signal is retransmitted to another earth station.



The time period for one complete orbital motion of an artificial satellite is equal to the time period of the earth's one complete rotation.

Advantages Of Satellite Microwave Communication:

- The coverage area of a satellite microwave is more than the terrestrial microwave.
- The transmission cost of the satellite is independent of the distance from the centre of the coverage area.
- Satellite communication is used in mobile and wireless communication applications.
- It is easy to install.
- It is used in a wide variety of applications such as weather forecasting, radio/TV signal broadcasting, mobile communication, etc.

Disadvantages Of Satellite Microwave Communication:

- Satellite designing and development requires more time and higher cost.
- The Satellite needs to be monitored and controlled on regular periods so that it remains in orbit.
- The life of the satellite is about 12-15 years. Due to this reason, another launch of the satellite has to be planned before it becomes non-functional.

Infrared Transmission

An infrared transmission is a wireless technology used for communication over short ranges. It is used for short-range communication such as data transfer between two cell phones, TV remote operation, data transfer between a computer and cell phone resides in the same closed area. Correspondingly, infrared frequencies are higher than those of microwaves, but lower than those of visible light.

Infrared wave lies in between visible light spectrum and microwaves. It has wavelength of 700-nm to 1-mm and frequency ranges from 300-GHz to 430-THz. Their wavelengths are shorter than microwaves but longer than visible light. Infrared propagation is line of sight.

They cannot penetrate walls due to its high frequency range and sun's infrared rays interfere with these rays. So cannot be used for long – range communication. As their usage is confined within closed space, they do not need any government permissions for their applications.

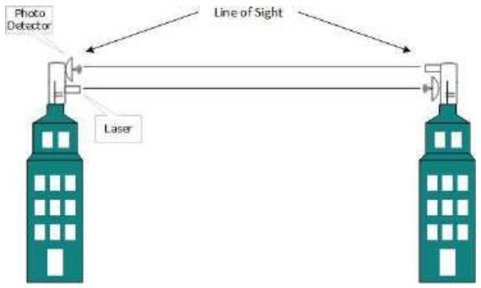
Characteristics Of Infrared:

- It supports high bandwidth, and hence the data rate will be very high.
- Infrared waves cannot penetrate the walls. Therefore, the infrared communication in one room cannot be interrupted by the nearby rooms.
- An infrared communication provides better security with minimum interference.
- Infrared communication is unreliable outside the building because the sun rays will interfere with the infrared waves

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.

2

Chapter

Data and Signals

LEARNING OBJECTIVES

In this chapter, we are going to discuss the following topics:

- ❑ Introduction
- ❑ Analog and Digital Data
- ❑ Analog and Digital Signals
- ❑ Signal Characteristics
- ❑ Digital Signal
- ❑ Comparison Between Analog and Digital Signal
- ❑ Digital Transmission
- ❑ Data Transmission Rate and the Bandwidth
- ❑ Performance Measurement of a Network

2.0. INTRODUCTION

Before we learn about analog and digital data, we must understand the concept of data. Generally data are entries that convey information. When two systems communicate with each other they send data in the form of signal through some media from one system to another. Here the term signal is defined as the electrical encoding (representation) of data. Let's take an example, when we talk in the telephone set, our speech generates sound waves, which generates electrical signals of a shape similar to the sound waves. These signals traverse across the telephone wires through various switches/exchanges and reach the telephone set of the receiver. At the receiver, the electrical signal generates sound waves of similar shape. Signalling is the act of propagation of signals through a suitable medium. Now understand the concept of data and signal and their significance in data communication.

Data: Data refers to information that conveys some meaning based on some mutually agreed up rules or conventions between a sender and a receiver and today it comes in a variety of forms such as text, graphics, audio, video and animation. Data can be categorized into two forms:

- Analog data
- Digital data

Analog data take on continuous values on some interval. Typical examples of analog data are voice and video. The data that are collected from the real world with the help of transducers are continuous-valued or analog in nature. On the contrary, digital data take on discrete values. Text or character strings can be considered as examples of digital data. Characters are represented by suitable codes, *e.g.*, ASCII code, where each character is represented by a 7-bit code.

Signal: It is electrical, electronic or optical representation of data, which can be sent over a communication medium. Stated in mathematical terms, a signal is merely a function of the data. For example, a microphone converts voice data into voice signal, which can be sent over a pair of wire. Analog signals are continuous-valued; digital signals are discrete-valued. The independent variable of the signal could be time (speech, for example), space (images), or the integers (denoting the sequencing of letters and numbers in the cricket score). The diagram below represents an analog signal.

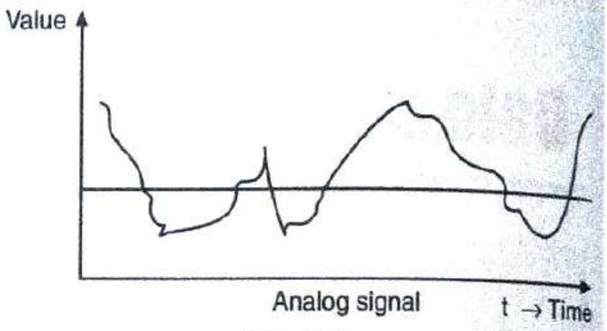


Fig. 2.1.

Digital signal can have only a limited number of defined values, usually two values 0 and 1, as shown in figure below.

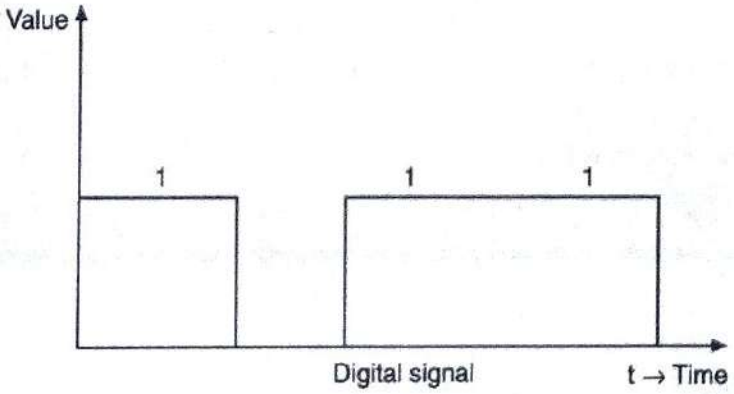


Fig. 2.2.

Signalling: It is an act of sending signal over communication medium.

Transmission: Communication of data by propagation and processing is known as transmission.

2.1. ANALOG AND DIGITAL DATA

The terms analog and digital correspond to continuous and discrete, respectively. Analog data takes on continuous values on some interval. The most familiar and common example of analog data is audio signal. Frequency components of speech may be found between 20 Hz and 20 kHz. The basic speech energy is concentrated between 300-3400 Hz. The frequencies up to 4000 Hz add very little to the intelligibility of human ear.

Another common example of analog data is video. The outputs of many sensors, such as temperature and pressure sensors, are also examples of analog data. Digital data takes on discrete values; *e.g.*, a computer's output.

- Analog transmission is a means of transmitting analog signals regardless of their content. The data may be analog or digital.
- Digital transmission is the transfer of information through a medium in digital form. A digital signal can be transmitted only for a limited distance.
- Data communications is the transfer of information that is in digital form, before it enters the communication system.

The two methods of sending data from computer *A* to computer *B*, both cases are examples of data communications, because the original data is digital in nature.

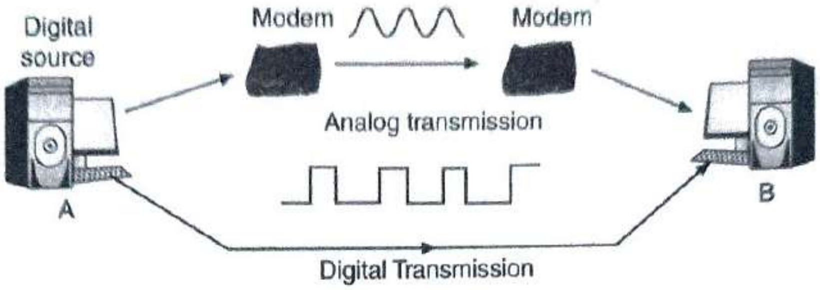
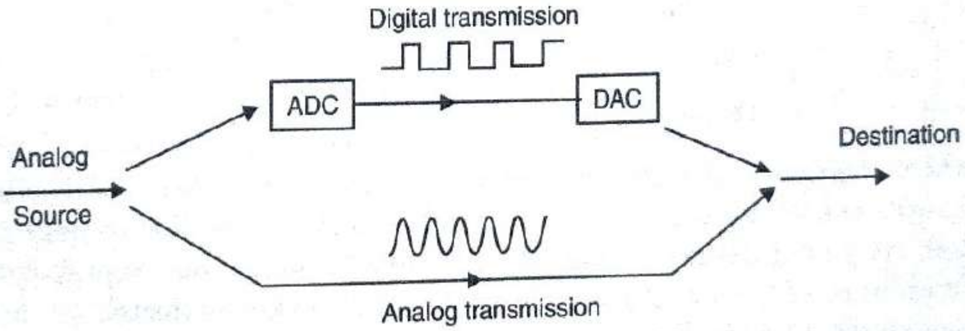


Fig. 2.3. Two methods of sending digital data

Two ways of transmitting analog information, in either cases it is not data communications, because the original information is not digital.



ADC: Analog-Digital-Converter
DAC: Digital-Analogt-Converter

Fig. 2.4. Two ways of transmitting analog information

2.2. ANALOG AND DIGITAL SIGNALS

Signal is generated by a transmitter and transmitted over a medium. Signal is a function of time or frequency. A signal is any function that carries information. Based on the range of variation of independent variables, signals can be divided into two classes: continuous-time (or analog) signals and discrete-time (or digital) signals. A signal is a function of time, but can also be expressed as a function of frequency; that is, the signal consists of components of different frequencies.

Any signal can be classified into one of the two types: analog and digital. An analog signal is a continuously varying signal, similar to a sinusoidal waveform. For instance, if we measure the room temperature continuously and plot a graph with the time on the *x*-axis and the temperature on the *y*-axis, we could get a continuous waveform, which is an analog signal. Analog is always continuous.

Analog signals can have any value in a range; while digital signals can have only a limited number of values.

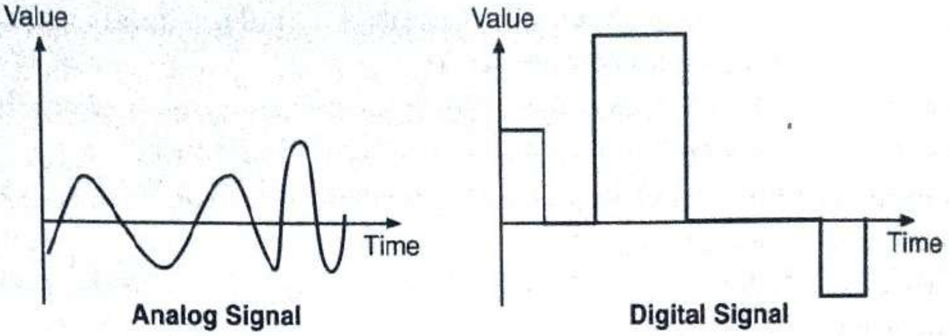


Fig. 2.5. Analog and digital signal waveform

The main characteristics of analog signals are:

- **Amplitude:** This is the strength of the signal. It can be expressed a number of different ways (as volts, decibels). The higher the amplitude, the stronger (louder) the signal. The decibel (named in honor of Alexander Graham Bell) is a popular measure of signal strength.

Sound level	Type of Sound
40 db	normal speech
90 db	lawn mowers
110 db	shotgun blast
120 db	jet engine taking off
120 db+	rock concerts

It has been discovered that exposure to sounds greater than 90 db for a period exceeding 15 minutes causes permanent damage to hearing. Our ability to hear high notes is affected. As young babies, we have the ability to hear quite high frequencies. This ability reduces as the aging process occurs. It can also be affected by too much noise over sustained periods. Ringing in the ears after being exposed to loud noise is an indication that hearing loss may be occurring.

This diagram shows a single signal of various amplitudes. The base line indicates a steady state, in this example; the signal amplitude rises both above and below the steady state. The measurement of the two extremes is called the peak to peak measurement.

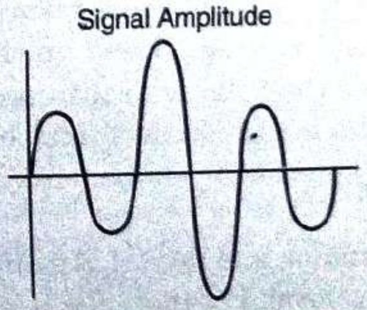


Fig. 2.6.

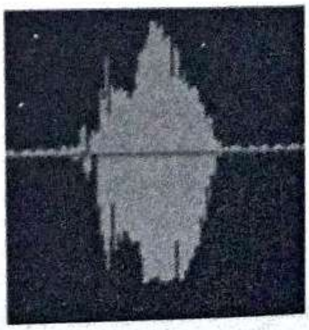


Fig. 2.7.

This Fig. 2.7, illustrates a speech signal, in this instance, the word "Hello". Speech is a very complex signal, and contains many thousands of different combinations of signals all mixed together. Note that it looks much more complicated than the single signal shown above in Fig. 2.7.

- **Frequency:** This is the rate of change the signal undergoes every second, expressed in Hertz (Hż), or cycles per second. A 30 Hz signal changes thirty times a second. In speech, we also refer to it as the number of vibrations per second. As we speak, the air is forced out of our mouths, being vibrated by our voice box. Men, on average, tend to vibrate the air at a lower rate than women, thus tend to have deeper voices.

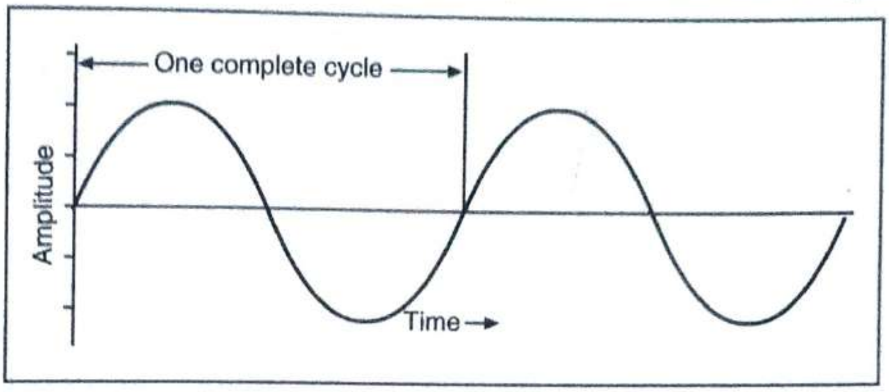


Fig. 2.8.

A cycle is one complete movement of the wave, from its original start position and back to the same point again. The number of cycles (or waves) within a one second time interval is called cycles-per-second, or Hertz.

- **Phase:** This is the rate at which the signal changes its relationship to time, expressed as degrees. One complete cycle of a wave begins at a certain point, and continues till the same point is reached again. Phase shift occurs when the cycle does not complete, and a new cycle begins before the previous one has fully completed.

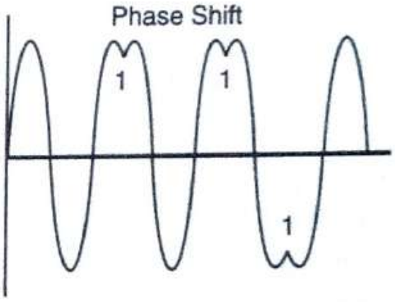


Fig. 2.9.

The human ear is insensitive to phase shift, but data signals are severely affected by it. Phase shift is caused by imperfections in cable media, such as joins and imperfect terminations. In a practical sense, imagine you are in the bath. If you drop the soap, it forms ripples where the waves travel outwards towards the edge of the bath. When the wave reaches the edge, it hits the wall of the bath and bounces back. This is like phase shift, which is an abrupt change in the signals relationship.

2.3. SIGNAL CHARACTERISTICS

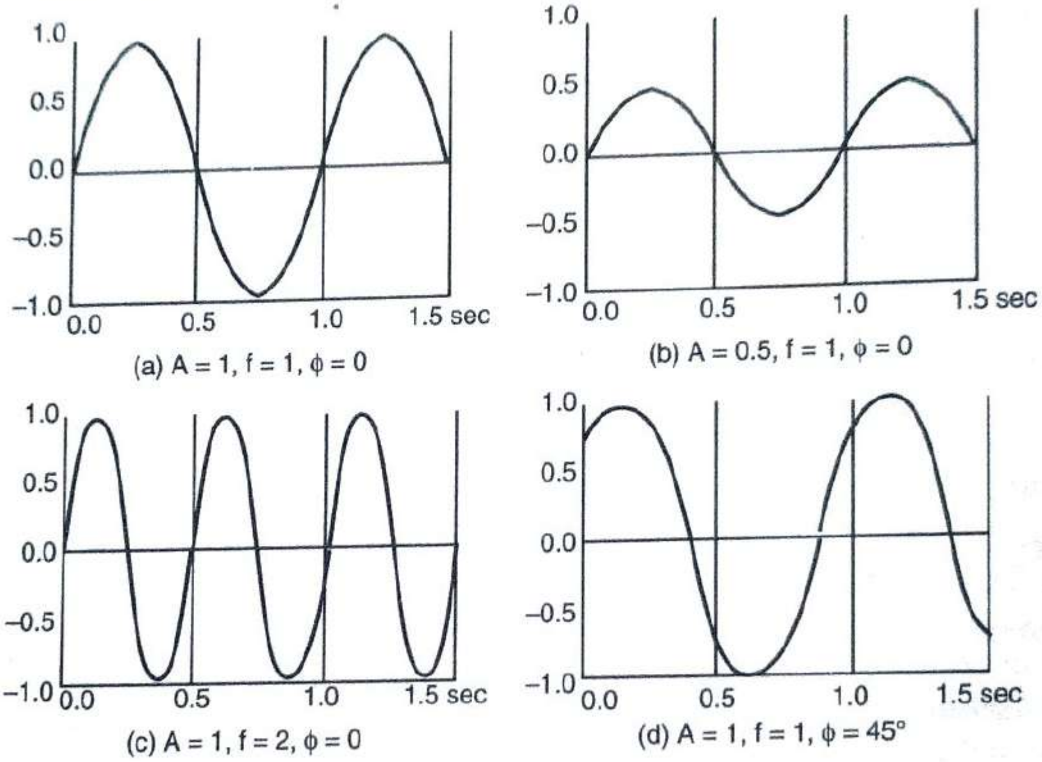
A signal can be represented as a function of time, *i.e.,* it varies with time. However, it can be also expressed as a function of frequency, *i.e.,* a signal can be considered as a composition of different frequency components. Thus, a signal has both time-domain and frequency domain representation.

2.3.1. Time-Domain Concepts

A signal is ***continuous*** over a period, if $\lim_{t \to a} s(t) = s(a)$, for all a, *i.e.,* there is no break in the signal. A signal is ***discrete*** if it takes on only a finite number of values. A signal is ***periodic*** if and only if $s(t + T) = s(t)$ for $-a < t < a$, where T is a constant, known as ***period***. The period is measured in seconds. In other words, a signal is a ***periodic signal*** if it completes a pattern within a measurable time frame. A periodic signal is characterized by the following three parameters.

- **Amplitude:** It is the value of the signal at different instants of time. It is measured in volts.
- **Frequency:** It is inverse of the time period, *i.e.*, $f = 1/T$. The unit of frequency is hertz (Hz) or cycles per second.
- **Phase:** It gives a measure of the relative position in time of two signals within a single period. It is represented by φ in degrees or radian.

A sine wave, the most fundamental periodic signal, can be completely characterized by its amplitude, frequency and phase. Examples of sine waves with different amplitude, frequency and phase are shown in Fig. 2.5. The phase angle φ indicated in the figure is with respect to the reference waveform shown in Fig. 2.5. (*a*).



(a) A = 1, f = 1, ϕ = 0

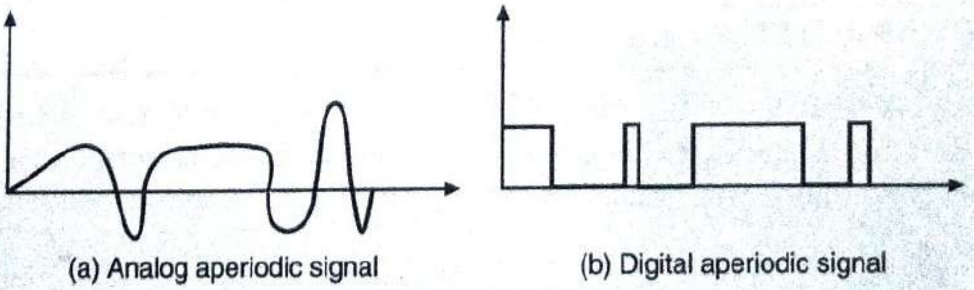
(b) A = 0.5, f = 1, ϕ = 0

(c) A = 1, f = 2, ϕ = 0

(d) A = 1, f = 1, ϕ = 45°

Fig. 2.10. Examples of signals with different amplitude, frequency and phase

An *aperiodic signal* or nonperiodic signal changes constantly without exhibiting a patter or cycle that repeats over time as shown in Fig. 2.11.



(a) Analog aperiodic signal

(b) Digital aperiodic signal

Fig. 2.11. Examples of aperiodic signals

2.3.2. Frequency Domain Concepts

The time domain representation displays a signal using *time-domain plot*, which show changes in signal amplitude with time. The time-domain plot can be visualized with the hel of an oscilloscope. The relationship between amplitude and frequency is provided by frequenc domain representation, which can be displayed with the help of *spectrum analyser*. Time domai and frequency domain representations of three sine waves of three different frequencies are shown in Fig. 2.12.

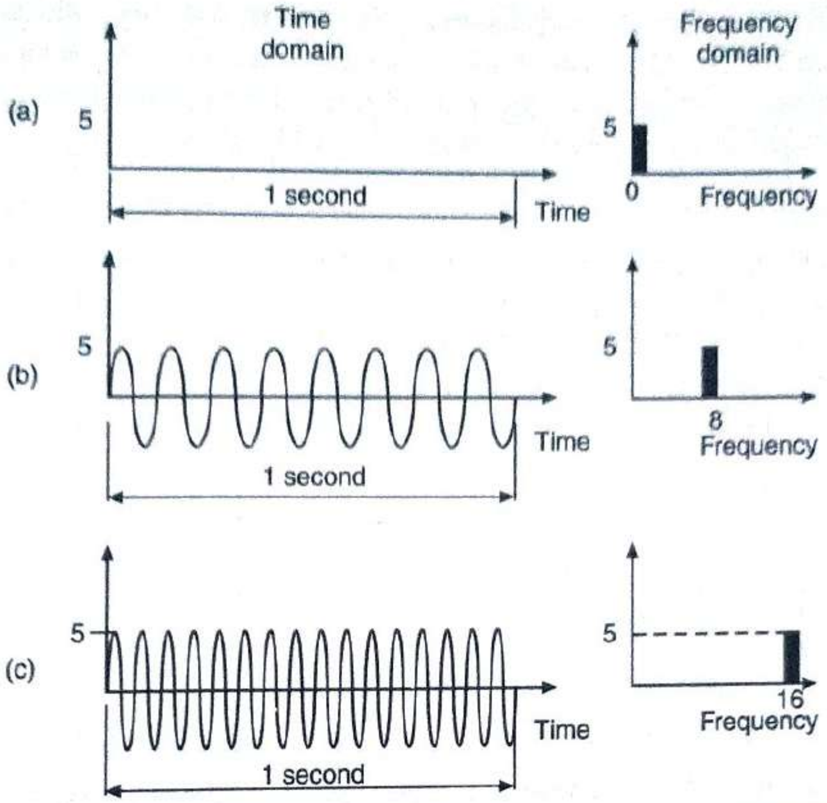
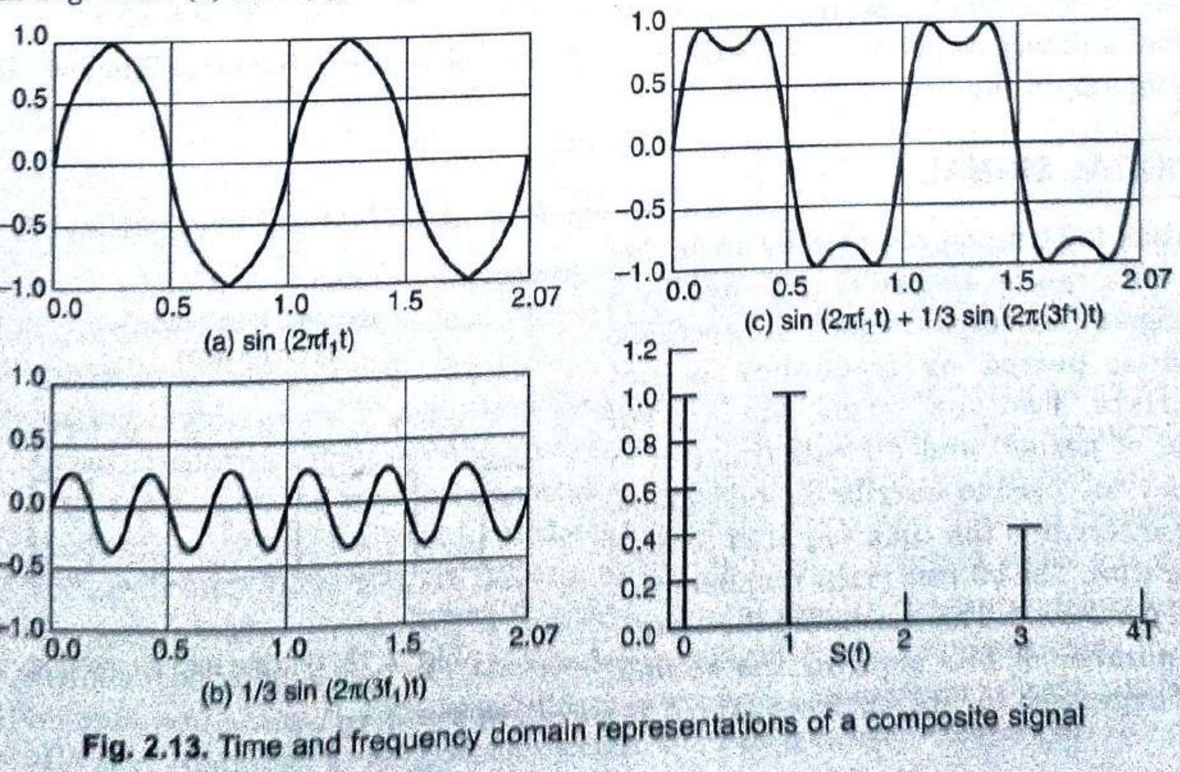


Fig. 2.12. Time domain and frequency domain representations of sine waves

Although simple sine waves help us to understand the difference between the time-domain and frequency domain representation, these are of little use in data communication. Composite signals made of many simple sine waves find use in data communication. Any composite signal can be represented by a combination of simple sine waves using Fourier analysis. For example, the signal shown in Fig. 2.13. (*c*) is a composition of two sine waves having frequencies f_1, $3f_1$, shown in Fig. 2.13. (*a*) and (*b*), respectively and it can be represented by



(a) sin $(2\pi f_1 t)$

(c) sin $(2\pi f_1 t)$ + 1/3 sin $(2\pi(3f_1)t)$

(b) 1/3 sin $(2\pi(3f_1)t)$

Fig. 2.13. Time and frequency domain representations of a composite signal

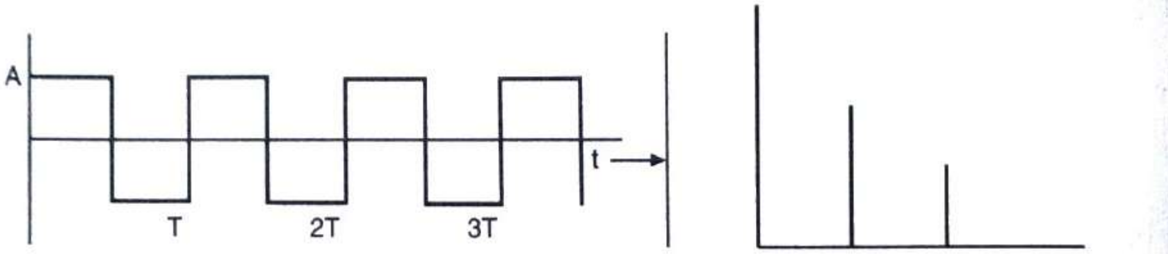
$$s(t) = \sin \omega t + 1/3 \sin 3\omega t\ ,$$

where $\omega = 2\pi f_1$.

The frequency domain function $s(f)$ specifies the constituent frequencies of the signal. The range of frequencies that a signal contains is known as it *spectrum*, which can be visualized with the help of a spectrum analyser. The band of frequencies over which most of the energy of a signal is concentrated is known as the *bandwidth* of the signal.

2.3.3. Frequency Spectrum

Frequency spectrum of a signal is the range of frequencies that a signal contains. *E.g.*, Consider a square wave shown in Fig. 2.14. (*a*). It can be represented by a series of sine waves $S(t) = 4A/\pi \sin 2\pi ft + 4A/3\pi \sin(2\pi(3f)t) + 4A/5\pi \sin 2\pi\ (5f)t + \ldots$ having frequency components f, $3f$, $5f$, ... and amplitudes $4A/\pi$, $4A/3\pi$, $4A/5\pi$ and so on. The frequency spectrum of this signal can be approximation comprising only the first and third harmonics as shown in Fig. 2.14. (*b*).



(a) A square wave (b) Frequency spectrum of a square wave

Fig. 2.14. (a) A square wave, (b) Frequency spectrum of a square wave

2.3.4. Bandwidth

The range of frequencies over which most of the signal energy of a signal is contained is known as **bandwidth** or effective bandwidth of the signal. The term 'most' is somewhat arbitrary. Usually, it is defined in terms of its 3dB cut-off frequency. The frequency spectrum and spectrum of a signal is shown in Fig. 2.10. Here the f_l and f_h may be represented by 3dB below ($A/\sqrt{2}$) the maximum amplitude.

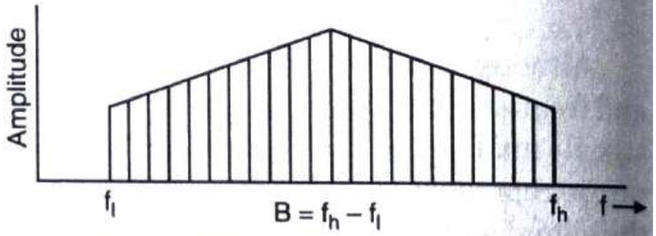


Fig. 2.15. Frequency spectrum and bandwidth of a signal

2.4. DIGITAL SIGNAL

In addition to being represented by an analog signal, data can be also be represented by a digital signal. Most digital signals are aperiodic and thus, period or frequency is not appropriate. Two new terms, *bit interval* (instead of period) and *bit rate* (instead of frequency) are used to describe digital signals. The bit interval is the time required to send one single bit. The bit rate is the number of bit interval per second. This mean that the bit rate is the number of bits send in one second, usually expressed in bits per second (bps) as shown in Fig. 2.16.

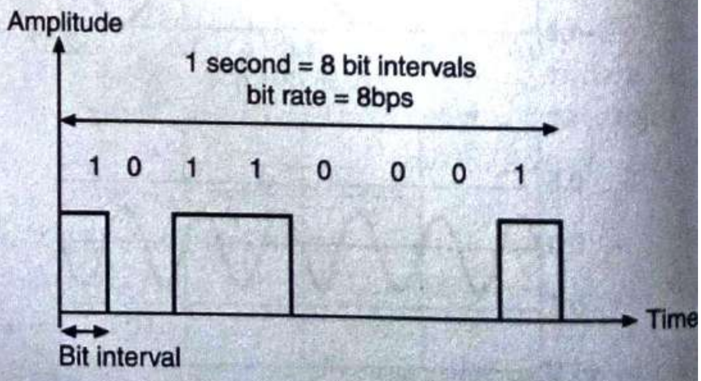


Fig. 2.16. Bit rate and bit interval

A digital signal can be considered as a signal with an infinite number of frequencies and transmission of digital requires a low-pass channel as shown in Fig. 2.17. On the other hand, transmission of analog signal requires band-pass channel shown in Fig. 2.18.

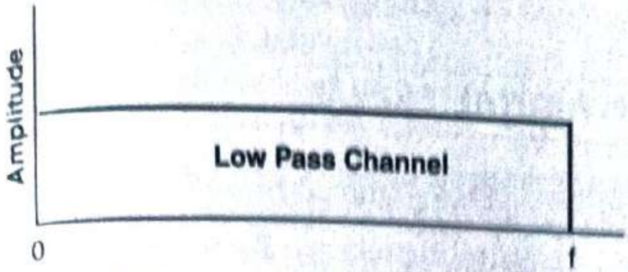


Fig. 2.17. Low pass channel required for transmission of digital signal

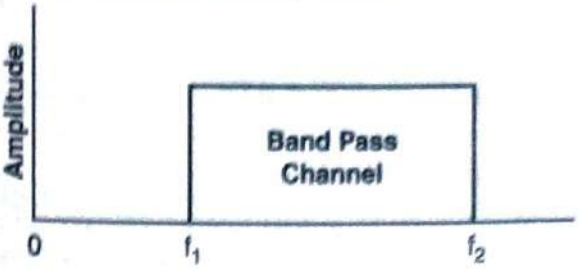
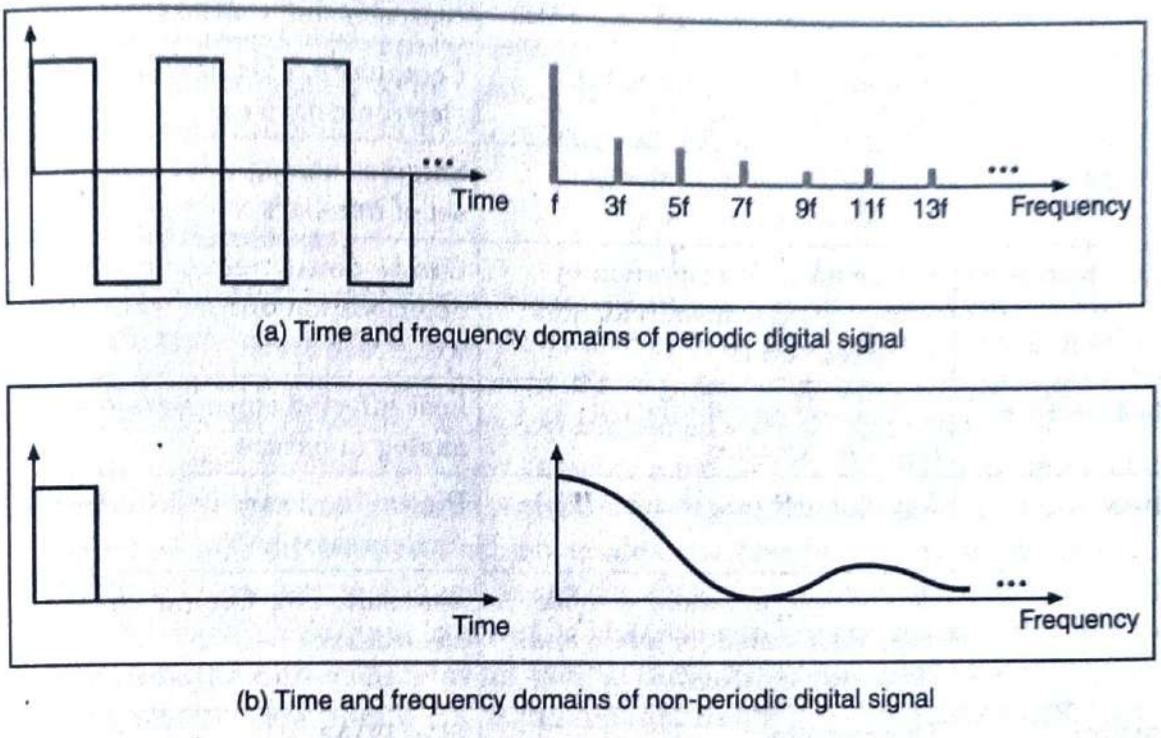


Fig. 2.18. Low pass channel required for transmission of analog signal

Digital Signal as a Composite Analog Signal: A digital signal is a composite analog signal with an infinite bandwidth.



(a) Time and frequency domains of periodic digital signal

(b) Time and frequency domains of non-periodic digital signal

Fig. 2.19.

2.4.1. Advantages of Digital Transmission

Digital transmission has several advantages over analog transmission. That is why there is a shift towards digital transmission despite large analog base. Some of the advantages of digital transmission are discussed below:

- Analog circuits require amplifiers, and each amplifier adds distortion and noise to the signal. In contrast, digital amplifiers regenerate an exact signal, eliminating cumulative errors. An incoming (analog) signal is sampled, its value is determined, and the node then generates a new signal from the bit value; the incoming signal is discarded. With analog circuits, intermediate nodes amplify the incoming signal, noise and all.
- Voice, data, video, etc. can all by carried by digital circuits. What about carrying digital signals over analog circuit? The modem example shows the difficulties in carrying digital over analog. A simple encoding method is to use constant voltage levels for a "1" and a "0". Can lead to long periods where the voltage does not change.

- Easier to multiplex large channel capacities with digital.
- Easy to apply encryption to digital data.
- Better integration if all signals are in one form. It can integrate voice, video and digital data.

2.5. COMPARISON BETWEEN ANALOG AND DIGITAL SIGNAL

Description	Analog Signal	Digital Signal
Signal	Analog signal is a continuous signal which represents physical measurements.	Digital signals are discrete time signals generated by digital modulation.
Waves	Denoted by sine waves.	Denoted by square waves.
Representation	Uses continuous range of values to represent information.	Uses discrete or discontinuous values to represent information.
Example	Human voice in air, analog electronic devices.	Computers, CDs, DVDs, and other digital electronic devices.
Technology	Analog technology records waveforms as they are.	Samples analog waveforms into a limited set of numbers and records them.
Data transmissions	Subjected to deterioration by noise during transmission and write/read cycle.	Can be noise-immune without deterioration during transmission and write/read cycle.
Response to Noise	More likely to get affected reducing accuracy.	Less affected since noise response are analog in nature
Flexibility	Analog hardware is not flexible.	Digital hardware is flexible in implementation.
Uses	Can be used in analog devices only. Best suited for audio and video transmission.	Best suited for Computing and digital electronics.
Applications	Thermometer.	PCs, PDAs.
Bandwidth	Analog signal processing can be done in real time and consumes less bandwidth.	There is no guarantee that digital signal processing can be done in real time and consumes more bandwidth to carry out the same information.
Memory	Stored in the form of wave signal.	Stored in the form of binary bit.
Power	Analog instrument draws large power.	Digital instrument draws only negligible power.
Cost	Low cost and portable.	Cost is high and not easily portable.
Impedance	Low.	High order of 100 megaohm.
Errors	Analog instruments usually have a scale which is cramped at lower end and give considerable observational errors.	Digital instruments are free from observational errors like parallax and approximation errors.

2.6. DIGITAL TRANSMISSION

Digital transmission is different from analog transmission, like digital signal is much simpler. Rather than being a continuously variable wave form, digital signal is a series of discrete pulses, which represents either a 1 bit or a 0 bit. Every computer uses a coding scheme that defines what combinations of 1s and 0s constitute all the characters in a character set.

How these 0s and 1s bits are carried through a network depends on whether the network is electrical or optical. In electrical networks, 1 bits are represented as high voltage and 0 bits are represented as absence of voltage or can be defined as null or low voltage. In optical networks, 1s bits are represented by the presence of light, and 0 bits are represented as the absence of light. The 1s and 0s, ON/OFF conditions are then carried through the network, and the receiving device repackages the 1s and 0s to determine what character is being represented. Digital signals are easier to reproduce than an analog signal, so it requires less care in the network. Besides using dumb amplifiers, a digital network uses regenerative repeaters. As a digital pulse travels over a distance, it may losses power, similar to an analog signal, such weakened and impaired signal enters the regenerative repeater, where the repeater examines the signal to determine what was supposed to be a one and what was supposed to be a 0. The repeater regenerates a new signal to pass on to the next point in the network, to eliminate noise and thus improving the error rate.

2.6.1. Baseband Transmission

Generally, a transmission signal contains more than a single frequency, either it may consist several different frequencies linked together or else superimposed on each other. *E.g.*, with today's communication technology it is virtually impossible to send low frequencies without experiencing any distortion. However, when low frequencies are sent the distortion is frequently so severe that the signal cannot even be used. So, what is the possible option to make low frequencies usable? The answer is that low frequency signals are copied and sent as higher frequencies for transmission purposes which makes the low frequency to reach their final destination. The low frequency would not reach the final destination without being distorted otherwise. These low frequency signals are called baseband signals and once they are changed to higher frequencies the name changes as well to radio frequency (RF) signals. A baseband signal has all the frequencies from 0 Hz to the highest frequency component with significant power. After the frequency is changed for transmission the higher frequency RF signal will have at least double what the baseband signal had initially. This simply means that the baseband signal is converted to the radio frequency in order to have twice the bandwidth. However, this bandwidth can also cause some transmission problems and undesirable results. So the radio frequency signal is subjected to filtering in order to reduce the bandwidth effect before transmitting. Once the RF signal is received it is demodulated and demultiplexed in order to achieve the original baseband signal.

As a result, defining baseband bandwidth is pretty easy. Simply put, baseband bandwidth is the bandwidth that exists before it is modulated or multiplexed or the baseband that results after demodulation and Demultiplexing. It is also noted that sometimes for short distance communications baseband signals are not modulated before being transmitted. This is because the transmission takes place over unmultiplexed lines which do not require complex modems or hardware in order to carry out the process. Good examples of this are the typical Ethernet and Token Ring that use these baseband signals.

2.6.2. Broadband Transmission

A signalling technology that sends signals simultaneously over a range of different frequencies as electromagnetic waves. The bandwidth of a broadband system can usually carry multiple simultaneous data signals. These signals are unidirectional-travelling in only one direction at a time. So a broadband system can generally ether transmit or receive but can not do both simultaneously. Broadband signals can be regenerate using amplifiers in order to travel longer distances before becoming attenuated.

How it works?

Broadband transmission is divided into multiple bands or channels by multiplexers using a multiplexing scheme such as Frequency division multiplexing (FDM). Each channel has a carrier frequency that is modulated to carry the signal from a given source. At the receiving station, multiplexers separate the various signals. Guard bands are used to prevent interference among channels.

Broadband transmission is typically used for environments in which video,.audio and data need to be transmitted simultaneously. Cable television systems are based on broadband transmission technologies, ATMs and DSL are some examples of broadband transmission.

2.6.3. Difference Between Baseband and Broadband Transmission

In Baseband, data is sent as digital signals through the media as a single channel tha uses the entire bandwidth of the media. Baseband communication is bi-directional, which mean that the same channel can be used to send and receive signals. In Baseband, frequency-divisio multiplexing is not possible. (Multiplexing (short muxing) is a process where multiple analo message signals or digital data streams are combined into one signal over a shared medium.

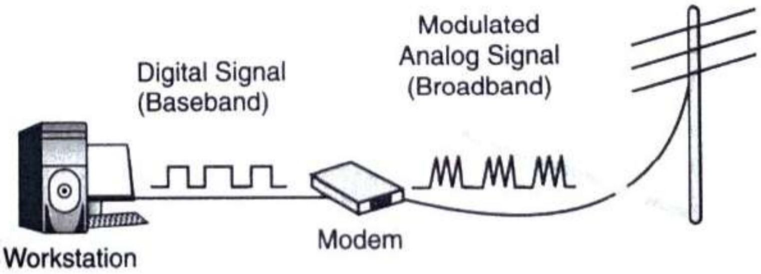


Fig. 2.20.

Broadband sends information in the form of an analog signal. Each transmission is assigne to a portion of the bandwidth; hence multiple transmissions are possible at the same tim Broadband communication is unidirectional, so in order to send and receive, two pathways a needed. This can be accomplished either by assigning a frequency for sending and assigning frequency for receiving along the same cable or by using two cables, one for sending and o for receiving. In broadband frequency-division multiplexing is possible.

2.7. DATA TRANSMISSION RATE AND THE BANDWIDTH

What is the connection between the bandwidth of a signal and the data rate that it represent Let's understand this by taking an example:

We can consider a sinusoidal wave of frequency = 10 Hz, *i.e.*, it completes one cycle 1/10th of a second, or it completes 10 cycles in one second.

We can imagine that the peak represents a binary 1 and the valley represents a binary By this way, this signal can also represent digital data. As each cycle represents two bits, t effective data rate is 10 × 2 = 20 bits per second or 20 bps. If we increase the frequency of the sinusoidal signal to 20 Hz, we will get a data rate of 40 bps and so on. Therefore, as the frequency becomes very high, the data rate that it represents also becomes very high.

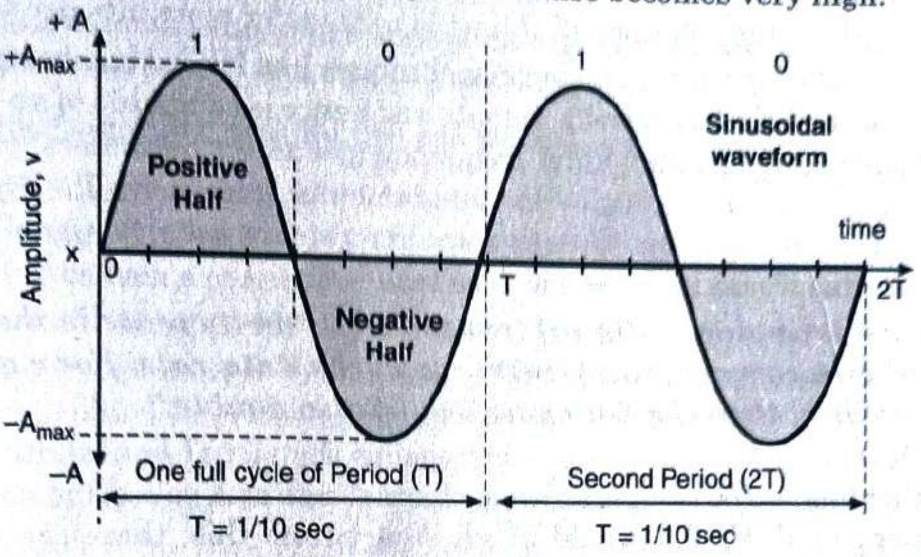


Fig. 2.21. A sinusoidal wave with frequency = 10 Hz

Representing digital data by a pure sinusoidal analog signal is dangerous. The equipment which has to measure the voltage to recognize a 1 and 0 will have to be super accurate and super sensitive. It has to measure the voltages exactly at the peaks and valleys. A small shift will result in a very large error and it will continue for all the subsequent readings also.

In a pure digital signal, the voltage levels corresponding to values 1 and 0 remain steady for a while for the reading to be taken more or less accurately. To get a digital signal would mathematically mean to get signals with frequencies between 0 to ∞ (infinity) bandwidth. This is not practical enough. There is not a medium available, which has an infinite bandwidth, which can carry any frequency. There has to be some alternative. Thus if we get a signal which is close to a digital signal it will also do.

Now consider following terms:

- **Signal Bandwidth:** The bandwidth of the transmitted signal or the range of frequencies present in the signal, as constrained by the transmitter, is called signal bandwidth.
- **Channel Bandwidth:** The range of signal bandwidths allowed by a communication channel without significant loss of energy (attenuation); is called channel bandwidth.
- **Channel Capacity or Maximum Data rate:** The maximum rate (in bps) at which data can be transmitted over a given communication link, or channel; is called channel capacity or maximum data rate. Channel capacity depends on four factors:
  1. Data rate (in bps).
  2. Bandwidth (constrained by transmitter and nature of transmission medium, expressed in cycles per second, or Hz).
  3. Noise (Average noise level over channel).
  4. Error rate (Percentage of time when bits are flipped).

In general, information is conveyed by change in values of the signal in time. Since frequency of a signal is a direct measure of the rate of change in values of the signal, the more the frequency of a signal, more is the achievable data rate or information transfer rate. This can be illustrated by taking the example of both an analog and a digital signal.

If we take analog transmission line coding techniques like Binary ASK, Binary FSK or Binary PSK, information is transferred by altering the property of a high frequency carrier wave. If we increase the frequency of this carrier wave to a higher value, then this reduces the bit interval T (= $1/f$) duration, thereby enabling us to transfer more bits per second.

Similarly, if we take digital transmission techniques like NRZ, Manchester encoding etc., these signals can be modelled as periodic signals and hence is composed of an infinite number of sinusoids, consisting of a fundamental frequency (f) and its harmonics. Here too, the bit interval (T) is equal to the reciprocal of the fundamental frequency ($T = 1/f$). Hence, if the fundamental frequency is increased, then this would represent a digital signal with shorter bit interval and hence this would increase the data rate.

"So, whether it is analog or digital transmission, an increase in the bandwidth of the signal, implies a corresponding increase in the data rate. For e.g., if we double the signal bandwidth, then the data rate would also double."

In practice however, we cannot keep increasing the signal bandwidth infinitely. The telecommunication link or the communication channel acts as a police and has limitations on the maximum bandwidth that it would allow. Apart from this, there are standardization transmission constraints that strictly limit the signal bandwidth to be used. So the achievable data rate is influenced more by the channel's bandwidth and noise characteristics than the signal bandwidth.

Nyquist and Shannon have given methods for calculating the channel capacity (C) of bandwidth limited communication channels.

2.7.1. Nyquist Criteria for Maximum Data Rate for Noiseless Channels

Given a noiseless channel with bandwidth B Hz., Nyquist stated that it can be used to carry atmost 2B signal changes (symbols) per second. The converse is also true, namely for achieving a signal transmission rate of 2B symbols per second over a channel, it is enough if the channel allows signals with frequencies upto B Hz.

Another implication of the above result is the sampling theorem, which states that for a signal whose maximum bandwidth is f Hz., it is enough to sample the signals at $2f$ samples per second for the purpose of quantization (A/D conversion) and also for reconstruction of the signal at the receiver (D/A conversion). This is because, even if the signals are sampled at a higher rate than $2f$ (and thereby including the higher harmonic components), the channel would anyway filter out those higher frequency components.

Also, symbols could have more than two different values, as is the case in line coding schemes like QAM, QPSK etc. In such cases, each symbol value could represent more than 1 digital bit.

Nyquist's formulae for multi-level signalling for a noiseless channel is

$$C = 2B \log_2 M$$

Where C is the channel capacity in bits per second, B is the maximum bandwidth allowed by the channel, M is the number of different signalling values or symbols and log is to the base 2.

For example, assume a noiseless 3-kHz channel.

- If binary signals are used, then $M = 2$ and hence maximum channel capacity or achievable data rate is $C = 2 * 3000 * \log 2 = 6000$ bps.
- Similarly, if QPSK is used instead of binary signalling, then M = 4. In that case, the maximum channel capacity is $C = 2 * 3000 * \log 4 = 2 * 3000 * 2 = 12000$ bps.

Thus, theoretically, by increasing the number of signalling values or symbols, we could keep on increasing the channel capacity C indefinitely. But however, in practice, no channel i noiseless and so we cannot simply keep increasing the number of symbols indefinitely, as the receiver would not be able to distinguish between different symbols in the presence of channel noise.

It is here that Shannon's theorem comes in handy, as he specifies a maximum theoretical limit for the channel capacity C of a noisy channel.

2.7.2. Shannon's Channel Capacity Criteria for Noisy Channels

Given a communication channel with bandwidth of B Hz. and a signal-to-noise ratio of S/N, where S is the signal power and N is the noise power, Shannon's formulae for the maximum channel capacity C of such a channel is

$$C = B \log (1 + S/N)$$

For example, for a channel with bandwidth of 3 KHz and with a S/N value of 30 DB, like that of a typical telephone line, the maximum channel capacity is

$$C = 3000 * \log (1 + 30) = 30000 \text{ bps (approx.)}$$

Using the previous examples of Nyquist criteria, we saw that for a channel with bandwidth 3 KHz, we could double the data rate from 6000 bps to 12000 bps, by using QPSK instead of binary signalling as the line encoding technique. Using Shannon's criteria for the same channel, we can conclude that irrespective of the line encoding technique used, we cannot increase the channel capacity of this channel beyond 30000 bps.

In practice however, due to receiver constraints and due to external noise sources, Shannon's theoretical limit is never achieved in practice.

Thus to summarize the relationship between bandwidth, data rate and channel capacity,

- In general, greater the signal bandwidth, the higher the information-carrying capacity
- But transmission system and receiver's capability limit the bandwidth that can be transmitted.

Hence data rate depends on:

- Available bandwidth for transmission.
- Channel capacity and Signal-to-Noise Ratio.
- Receiver Capability.

More the frequency allotted, more the channel bandwidth, more the processing capability of the receiver, greater the information transfer rate that can be achieved.

2.8. PERFORMANCE MEASUREMENT OF A NETWORK

One important issue in networking is the performance of the network—how good is it? Here we are going to discuss some terms which play an important role to determine the overall performance of a network. These terms are:

- Bandwidth
- Throughput
- Latency (delay)
- Bandwidth-delay product
- Jitter

1. **Bandwidth:** In networking, we use the term bandwidth in two contexts.
    - The first, bandwidth in hertz, refers to the range of frequencies in a composite signal or the range of frequencies that a channel can pass.

- The second, bandwidth in bits per second (bps), refers to the speed of bit transmission in a channel or link.

The bandwidth of a subscriber line is 4 kHz for voice or data. The bandwidth of this line for data transmission can be up to 56,000 bps using a sophisticated modem to change the digital signal to analog. If the telephone company improves the quality of the line and increases the bandwidth to 8 kHz, we can send 112,000 bps by using the same technology.

2. **Throughput:** The throughput is a measure of how fast we can actually send data through a network. For example, a network with bandwidth of 10 Mbps can pass only an average of 12,000 frames per minute with each frame carrying an average of 10,000 bits. We can calculate the throughput as:

$$\text{Throughput} = \frac{12{,}000 \times 10{,}000}{60} = 2 \text{ Mbps}$$

The throughput is almost one-fifth of the bandwidth in this case.

3. **Latency (Delay):** How long it takes for an entire message to completely arrive at the destination from the time the first bit is sent out from the source, is called *Latency.*

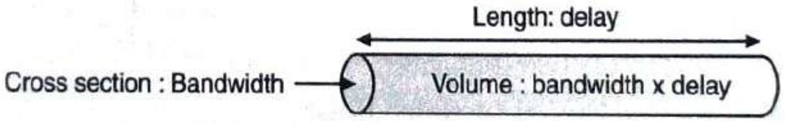
Latency = Propagation Time + Transmission Time + Queuing Time + Processing Time

For example; what is the propagation time if the distance between the two points is 12,000 km? Assume the propagation speed to be 2.4×10^8 m/s in cable.

$$\text{Propagation Time} = \frac{12{,}000 \times 1000}{2.4 \times 10^8} = 50 \text{ ms.}$$

4. **Bandwidth-Delay Product:** The bandwidth-delay product defines the number of bits that can fill the link.

$$\text{Bandwidth} - \text{Delay Product} = \text{Bandwidth} \times \text{Delay}$$



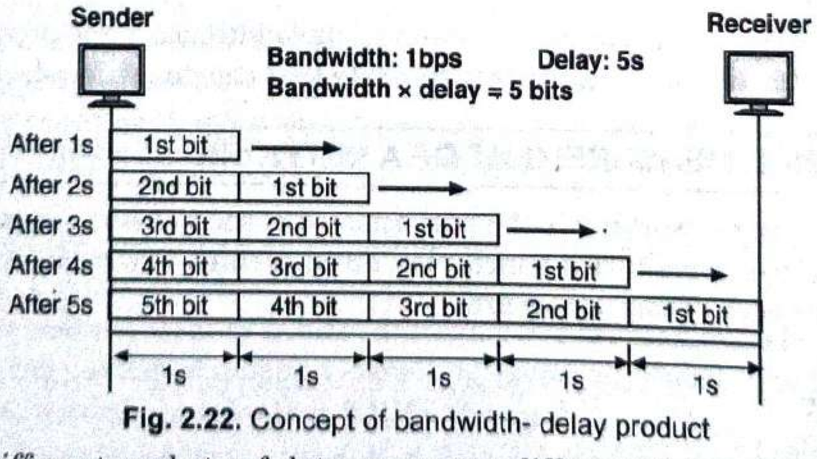


Fig. 2.22. Concept of bandwidth- delay product

5. **Jitter:** Different packets of data encounter different delays. It is a problem if the application using the data at the receiver site is time-sensitive.

Points to Remember

1. When two systems communicate with each other they send data in the form of signal through some media from one system to another.

1. Signals are defined as the electrical encoding (representation) of data.
2. The terms analog and digital correspond to continuous and discrete, respectively.
3. Analog data takes on continuous values on some interval. The most familiar and common example of analog data is audio signal.
4. Analog transmission is a means of transmitting analog signals regardless of their content. The data may be analog or digital.
5. Digital transmission is the transfer of information through a medium in digital form. A digital signal can be transmitted only for a limited distance.
6. Data communications is the transfer of information that is in digital form, before it enters the communication system.
7. Signal is a function of time or frequency. A signal is any function that carries information.
8. Amplitude is the strength of the signal. It can be expressed a number of different ways (as volts, decibels). The higher the amplitude, the stronger (louder) the signal.
9. Frequency is the rate of change the signal undergoes every second, expressed in Hertz (Hz), or cycles per second.
10. A cycle is one complete movement of the wave, from its original start position and back to the same point again.
11. Phase is the rate at which the signal changes its relationship to time, expressed as degrees.
12. A signal is ***continuous*** over a period, if $\lim_{t \to a} s(t) = s(a)$, for all a, *i.e.*, there is no break in the signal.
13. A signal is ***discrete*** if it takes on only a finite number of values.
14. A signal is ***periodic*** if and only if $s\,(t + T) = s(t)$ for $-a < t < a$, where T is a constant, known as ***period***. The period is measured in seconds. In other words, a signal is a ***periodic signal*** if it completes a pattern within a measurable time frame.
15. The range of frequencies over which most of the signal energy of a signal is contained is known as **bandwidth** or effective bandwidth of the signal.
16. The bit interval is the time required to send one single bit.
17. The bit rate is the number of bit interval per second. This mean that the bit rate is the number of bits send in one second, usually expressed in bits per second (bps).
18. *Base-band* is defined as one that uses digital signalling, which is inserted in the transmission channel as voltage pulses.
19. On the other hand, *broadband* systems are those, which use analog signalling to transmit information using a carrier of high frequency.

Solved Questions

Q.1. Distinguish between data and signal.

Ans. Data is an entity, which conveys some meaning. On the other hand, the signal is a representation of data in some electric, electromagnetic or optical form. So, whenever data needs to be sent, it has to be converted into signal of some form for transmission over a suitable medium.

Q.2. What do you mean by a "Periodic Signal"? And what are the three parameters that characterize it?

Ans. *A signal is periodic signal* if it completes a pattern within a measurable timeframe. A periodic signal is characterized by the following three parameters.

- **Amplitude:** It is the value of the signal at different instants of time. It is measured in volts.
- **Frequency:** It is inverse of the time period, *i.e.*, $f = 1/T$. The unit of frequency is Hertz (Hz) or cycles per second.

- **Phase:** It gives a measure of the relative position in time of two signals within a single period.

Q.3. Distinguish between time domain and frequency domain representation of a signal.

Ans. Whenever a signal is represented as a function of time, it is called time domain **representation**. An electromagnetic signal can be either continuous or discrete. It is represented as **$s(t)$**.

Whenever a signal is represented as a function of frequency, it is called **frequency domain** representation. It is expressed in terms of different frequency components and **represented** as $s(f)$.

Q.4. What equipments are used to visualize electrical signals in time domain and frequency domain?

Ans. Cathode Ray Oscilloscope is used to visualize electrical signals in time domain and Spectrum Analyser used to visualize electrical signals in frequency domain.

Q.5. What do you mean by the Bit Interval and Bit rate in a digital signal?

Ans. The bit interval is the time required to send one single bit. The bit rate is the number of bit intervals per second. This mean that the bit rate is the number of bits send in one second, usually expressed in bits per second (bps).

Q.6. Define baseband.

Ans. In Baseband, data is sent as digital signals through the media as a single channel that uses the entire bandwidth of the media. Baseband communication is bi-directional, which means that the same channel can be used to send and receive signals. In Baseband, frequency-division multiplexing is not possible. (Multiplexing (short muxing) is a process where multiple analog message signals or digital data streams are combined into one signal over a shared medium.)

Q.7. Define broadband.

Ans. Broadband sends information in the form of an analog signal. Each transmission is assigned to a portion of the bandwidth; hence multiple transmissions are possible at the same time. Broadband communication is unidirectional, so in order to send and receive, two pathways are needed. This can be accomplished either by assigning a frequency for sending and assigning a frequency for receiving along the same cable or by using two cables, one for sending and one for receiving. In broadband frequency-division multiplexing is possible.

Exercise

A. Multiple Choice Questions

1. If the maximum amplitude of a sine wave is 2 V, the minimum amplitude is .... V.
   (*a*) 2 (*b*) 1 (*c*) – 2 (*d*) Between – 2 and 2

2. .... can impair a signal.
   (*a*) Noise (*b*) Attenuation (*c*) Distortion (*d*) All of the above

3. .... is the rate of change with respect to time.
   (*a*) Time (*b*) Frequency (*c*) Amplitude (*d*) Voltage

4. A signal is measured at two different points. The power is P_1 at the first point and P_2 at the second point. The dB is 0. This means ....
   (*a*) P_2 equals P_1 (*b*) P_2 is zero
   (*c*) P_2 is much larger than P_1 (*D*) P_2 is much smaller than P_1

5. Baseband transmission of a digital signal is possible only if we have a .... channel.
   (*a*) Bandpass (*b*) Low-pass (*c*) High rate (*d*) Low rate

6. .... is a type of transmission impairment in which the signal loses strength due to the resistance of the transmission medium.
(*a*) Distortion (*b*) Attenuation (*c*) Noise (*d*) Decibel

7. A sine wave in the .... domain can be represented by one single spike in the .... domain.
(*a*) Time; phase (*b*) Frequency; time (*c*) Time; frequency (*d*) Phase; time

8. If the bandwidth of a signal is 5 KHz and the lowest frequency is 52 KHz, what is the highest frequency?
(*a*) 5 KHz (*b*) 47 KHz (*c*) 57 KHz (*d*) 10 KHz

9. In a time-domain plot, the horizontal axis is a measure of ....
(*a*) Phase (*b*) Signal amplitude (*c*) Frequency (*d*) Time

10. .... data are continuous and take continuous values.
(*a*) Digital (*b*) Analog (*c*) (*a*) or (*b*) (*d*) None of the above

11. Frequency and period are ....
(*a*) Proportional to each other (*b*) Inverse of each other
(*c*) The same (*d*) None of the above

12. When propagation speed is multiplied by propagation time, we get the ....
(*a*) Wavelength of the signal (*b*) Throughput
(*c*) Distance a signal or bit has traveled (*d*) Distortion factor

13. A .... sine wave is not useful in data communications; we need to send a .... signal.
(*a*) Single-frequency; composite (*b*) Composite; single-frequency
(*c*) Single-frequency; double-frequency (*d*) None of the above

14. The .... product defines the number of bits that can fill the link.
(*a*) Delay-amplitude (*b*) Frequency-amplitude
(*c*) Bandwidth-period (*d*) Bandwidth-delay

15. .... signals can have only a limited number of values.
(*a*) Digital (*b*) Analog (*c*) (*a*) or (*b*) (*d*) None of the above

16. Before data can be transmitted, they must be transformed to ....
(*a*) Periodic signals (*b*) Electromagnetic signals
(*c*) Aperiodic signals (*d*) Low-frequency sine waves

17. Data can be ....
(*a*) Digital (*b*) Analog (*c*) (*a*) or (*b*) (*d*) None of the above

18. .... is a type of transmission impairment in which the signal loses strength due to the different propagation speeds of each frequency that makes up the signal.
(*a*) Noise (*b*) Distortion (*c*) Attenuation (*d*) Decibel

19. Signals can be ....
(*a*) Digital (*b*) Analog (*c*) Either (*a*) or (*b*) (*d*) Neither (*a*) nor (*b*)

20. A sine wave is ....
(*a*) Periodic and discrete (*b*) Aperiodic and discrete
(*c*) Periodic and continuous (*d*) Aperiodic and continuous

21. .... data have discrete states and take discrete values.
(*a*) Analog (*b*) Digital (*c*) (*a*) or (*b*) (*d*) None of the above

22. For a .... channel, we need to use the Shannon capacity to find the maximum bit rate.
(*a*) Noiseless (*b*) Noisy (*c*) Low-pass (*d*) Bandpass

23. What is the bandwidth of a signal that ranges from 1 MHz to 4 MHz?
(*a*) 1 KHz (*b*) 3 MHz (*c*) 4 MHz (*d*) None of the above

24. .... signals can have an infinite number of values in a range.
(*a*) Analog (*b*) Digital (*c*) (*a*) or (*b*) (*d*) None of the above

25. A (*n*) .... signal is a composite analog signal with an infinite bandwidth.
(*a*) Digital (*b*) Analog (*c*) Either (*a*) or (*b*) (*d*) Neither (*a*) nor (*b*)

26. A periodic signal completes one cycle in 0.001 s. What is the frequency?
(*a*) 1 Hz (*b*) 100 Hz (*c*) 1 KHz (*d*) 1 MHz

27. The .... of a composite signal is the difference between the highest and the lowest frequencies contained in that signal.
(*a*) Period (*b*) Bandwidth (*c*) Frequency (*d*) Amplitude

28. .... is a type of transmission impairment in which an outside source such as crosstalk corrupts a signal.
(*a*) Noise (*b*) Distortion (*c*) Attenuation (*d*) Decibel

29. .... describes the position of the waveform relative to time 0.
(*a*) Amplitude (*b*) Phase (*c*) Frequency (*d*) Voltage

30. Given two sine waves *A* and *B*, if the frequency of *A* is twice that of *B*, then the period of *B* is .... that of *A*.
(*a*) One-half (*b*) Twice (*c*) The same as (*d*) Indeterminate from

31. As frequency increases, the period ....
(*a*) Increases (*b*) Decreases (*c*) Doubles(*d*) Remains the same

32. If the available channel is a .... channel, we cannot send a digital signal directly to the channel.
(*a*) Low-pass (*b*) Low rate (*c*) Bandpass (*d*) High rate

33. For a .... channel, the Nyquist bit rate formula defines the theoretical maximum bit rate.
(*a*) Low-pass (*b*) Bandpass (*c*) Noisy (*d*) Noiseless

34. In a frequency-domain plot, the horizontal axis measures the ....
(*a*) Phase (*b*) Frequency (*c*) Slope (*d*) Peak amplitude

Answers

1. (*c*); 2. (*d*); 3. (*b*); 4. (*a*); 5. (*b*); 6. (*b*); 7. (*c*); 8. (*c*); 9. (*d*); 10. (*b*); 11. (*b*); 12. (*c*); 13. (*a*); 14. (*d*); 15. (*a*); 16. (*b*); 17. (*c*); 18. (*b*); 19. (*c*); 20. (*c*); 21. (*b*); 22. (*b*); 23. (*b*); 24. (*a*); 25. (*a*); 26. (*c*); 27. (*b*); 28. (*a*); 29. (*b*); 30. (*b*); 31. (*b*); 32. (*c*); 33. (*d*); 34. (*b*).

B. Fill in the Blanks

1. A signal is a .... or Electromagnetic or .... data.
2. The four parameters that are used to characterize a signal in time domain concept are amplitude, ..., ... and ....
3. With the help of ...., we can find out the different frequency components of a signal, and these components are visualized with the help of ....
4. .... gives a measure of the span of the spectral components of a signal.
5. The noise created by the agitation of electrons of the transmission channel is termed as ....
6. The noise created by bunching several cables together is known as ....

Answers

1. electric, representation of data; 2. frequency phase, period; 3. spectrum, oscilloscope spectrum analyzer; 4. Spectrum analyser; 5. white noise; 6. crosstalk.

Review Questions

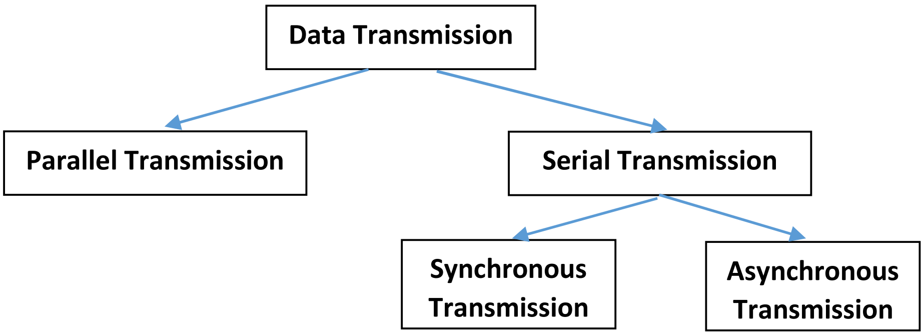
1. Define the term signal. What do you mean by analog and digital signals?
2. Explain the term bandwidth. Why it is useful?
3. Discuss the terms amplitude, period, frequency and phase.
4. Explain baseband and broadband signals.
5. Explain various signals characteristics.
6. What are the differences between analog and digital signals?
7. Explain digital transmission.
8. Define data transmission rate.
9. Explain the Nyquist and Shannon capacities.
10. What are the various factors which defines the performance of a network?

UNIT - 4
Data Communication & Data link control

4.1 Asynchronous and Synchronous Transmission
4.2 Error Detection
4.3 Line configuration
4.4 Flow Control
4.5 Error Control
4.6 Multiplexing
4.7 FDM
4.8 Synchronous TDM and Asynchronous (Statistical) TDM

4.1 DATA TRANSMISSION :

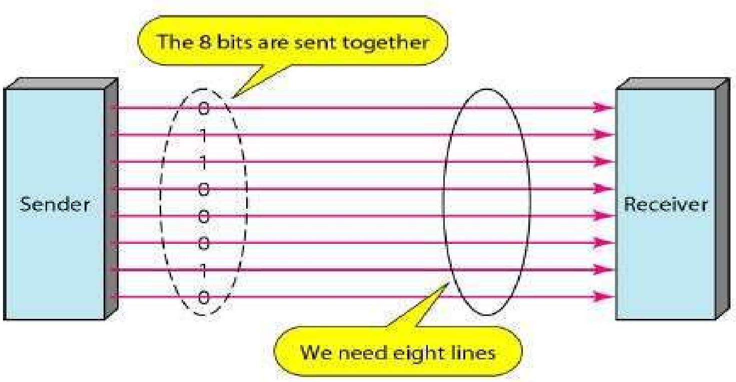
Data is transmitted from one device to another in analog or digital format. There are two methods to transmit data between two digital devices. When we transmit data from device to another, data can be transmitted in parallel mode or serial mode. In parallel mode, multiple bits are sent with each clock pulse. But in serial mode, 1 bit is send with each clock pulse. There is only one way to send parallel transmission whereas there are two sub-ways to send serial transmission.


Data Transmission
Parallel Transmission
Serial Transmission
Synchronous Transmission
Asynchronous Transmission

PARALLEL TRANSMISSION :

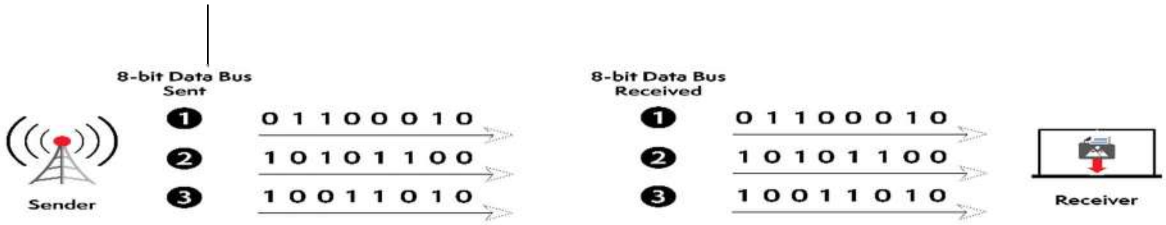
Binary data consisting of 1s and 0s. Binary data may be organised into groups of n-bits each. By grouping, we can send data n-bits at a time instead of a single bit. This type of transmission is called parallel transmission.

When data is sent using parallel data transmission, multiple data bits are transmitted over multiple channels at the same time. This means that data can be sent much faster than using serial transmission methods.

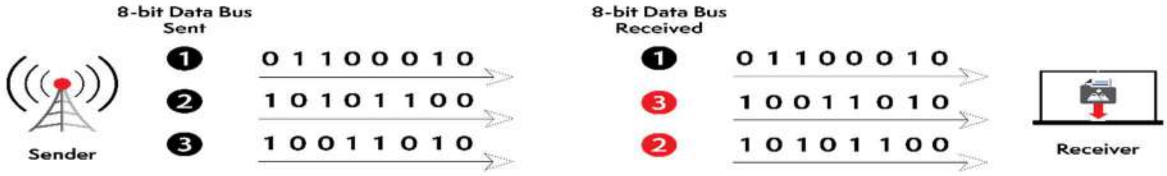


Given that multiple bits are sent over multiple channels at the same time, the order in which a bit string is received can depend on various conditions, such as proximity to the data source, user location, and bandwidth availability.

Examples :



The data is sent and received in the correct order.



The data is sent in the correct order, but some bits were received faster than others.

Advantages :

The main advantages of parallel transmission over serial transmission are:

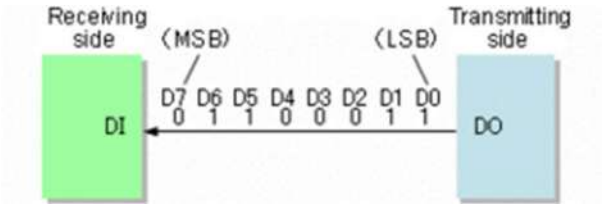
- it is easier to program;
- and data is sent faster.

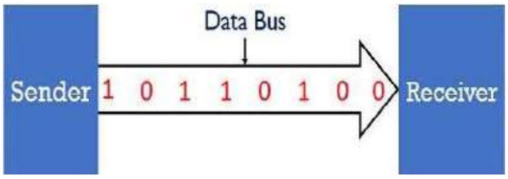
Dis-advantages:

- Requires more transmission channels than serial transmission
- Data bits can be out of sync, depending on transfer distance and how fast each bit loads.
- Expensive

SERIAL TRANSMISSION :

In serial transmission, data can be sent bit by bit from one computer to another in two directions. Each bit has a clock pulse rate. Eight bits are transmitted at a time, with a start and stop bit known as a parity bit, which is 0 and 1, respectively. So, we need only one communication channel rather than n-bit. Data cables are used when transmitting data over a longer distance. The data cable has D-shaped 9 pin cable that connects the data in series.





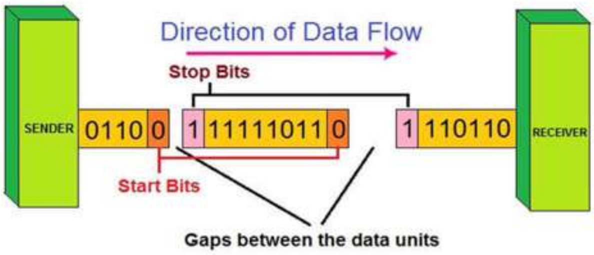
Categories of Serial Transmission :

Asynchronous transmission – an extra bit is added to each byte to alert the receiver to the arrival of new data. 0 is used as a start bit, while 1 used as a stop bit.

In asynchronous transmission, the timing of a signal is not important. The receiving device can retrieve the information without regard to the sequence in which it is sent. It is based on grouping the bit stream into bytes. Each group consists of 8 bits and sent along the link as a unit. The sending system handles each group independently.

The receiver can not use timing to predict when next group will arrive. To alert the receiver about arrival of a new group, an extra bit is added at the beginning of each byte. This bit is called Start Bit and it is usually 0. Similarly, to inform the receiver about finish of the byte, one more byte is added at the end of the byte. This is called Stop Bit and it is usually 1.

Asynchronous Communication



By this method, each byte is increased in size to 10 bits, out of which 8bits are data bits and 2 bits are signals (Start / Stop bits) to receiver. The transmission of each byte may be followed by a gap of varying duration by an idle channel or by a stream of additional stop bits.

The start & stop bits and the gap alert the receiver to beginning and end of each byte and allow it to synchronise with the data stream. This mechanism is called asynchronous transmission. In this method, we send one start bit (0) at the beginning and one stop bit (1) at the end of each byte.

Characteristics of Asynchronous Transmission

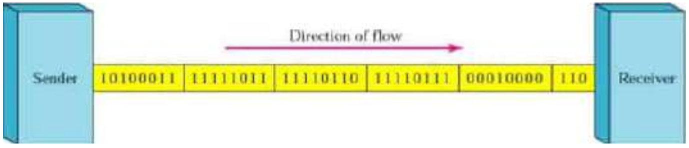
- Each character is headed by a beginning bit and concluded with one or more end bits.
- There may be gaps or spaces in between characters.

Examples of Asynchronous Transmission

- Emails
- Forums
- Letters
- Radios
- Televisions

Synchronous transmission – no extra bit is added to each byte. Data is transferred in batches, each of which contains multiple bytes.

In synchronous transmission, the bit stream is combined into longer frames which contains multiple bytes. Each byte is introduced onto the transmission link without gap. Data are transmitted as an unbroken string of 1s and 0s. it is left to the receiver to separate the bit stream into bytes for decoding purpose. Data are transmitted as an unbroken string of 0s and 1s. The receiver separates the string into bytes / characters to reconstruct the information.



In this method, we send bits one-after-another without any start/stop bits. It is the responsibility of the receive to group the bits. Synchronous transmission is effective, dependable, and often utilised for transmitting a large amount of data. It offers real-time communication between linked devices.

Synchronisation between the source and target is required so that the source knows where the new byte begins, since there are no spaces included between the data.

Because there are no beginning and end bits, the data transfer rate is quicker but there's an increased possibility of errors occurring. Some bytes could become damaged on account of lost bits. To resolve this issue, it's necessary to regularly re-synchronise the clocks, as well as to make use of check digits to ensure that the bytes are correctly received and translated.

Characteristics of Synchronous Transmission

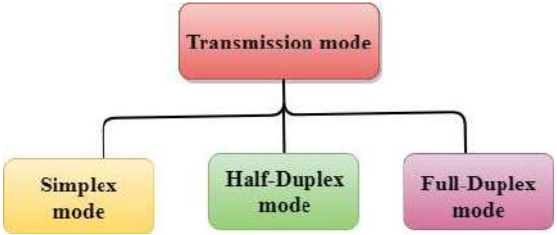
- There are no spaces in between characters being sent.
- Timing is provided by modems or other devices at the end of the transmission.
- Special 'syn' characters goes before the data being sent.
- The syn characters are included between chunks of data for timing functions.

Examples of Synchronous Transmission

- Chatrooms
- Video conferencing
- Telephonic conversations
- Face-to-face interactions

DATA TRANSMISSION MODE :

The transmission mode defines the direction of signal flow between two connected devices. There are three modes of transmission, namely: simplex, half duplex, and full duplex.

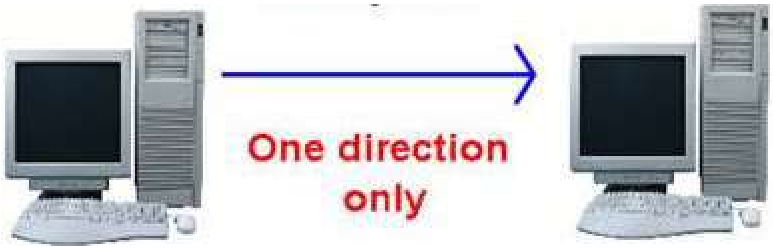


The primary difference between three modes of transmission is that in a simplex mode of transmission the communication is unidirectional, or one-way; whereas in the half duplex mode of transmission the communication is two-directional, but the channel is interchangeably used by both of the connected devices. On the other hand, in the full duplex mode of transmission, the communication is bi-directional or two-way, and the channel is used by both of the connected devices simultaneously.

Simplex

In simplex transmission mode, the communication between sender and receiver occurs in only one direction. The sender can only send the data, and the receiver can only receive the data. The receiver cannot reply to the sender.

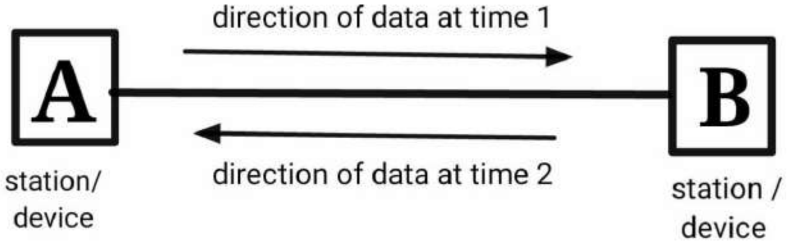
Simplex transmission can be thought of as a one-way road in which the traffic travels only in one direction—no vehicle coming from the opposite direction is allowed to drive through.



Example : keyboard / monitor relationship with CPU, TV programs etc

Half Duplex

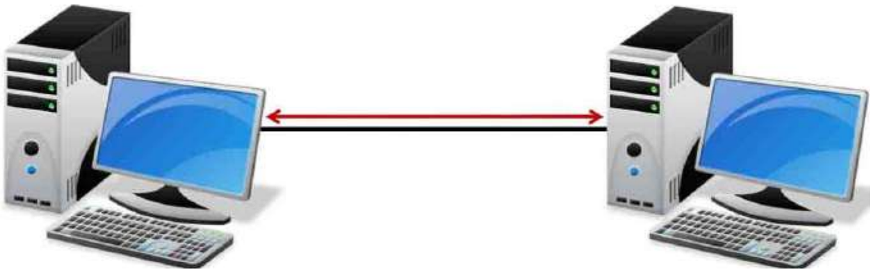
The communication between sender and receiver occurs in both directions in half duplex transmission, but only one at a time. The sender and receiver can both send and receive the information, but only one is allowed to send at any given time. When one devices transmits some data, the other devices can only receive. Half duplex is still considered a one-way road, in which a vehicle traveling in the opposite direction of the traffic has to wait till the road is empty before it can pass through.



Example : walkie-talkies, the speakers at both ends can speak, but they have to speak one by one. They cannot speak simultaneously.

Full Duplex

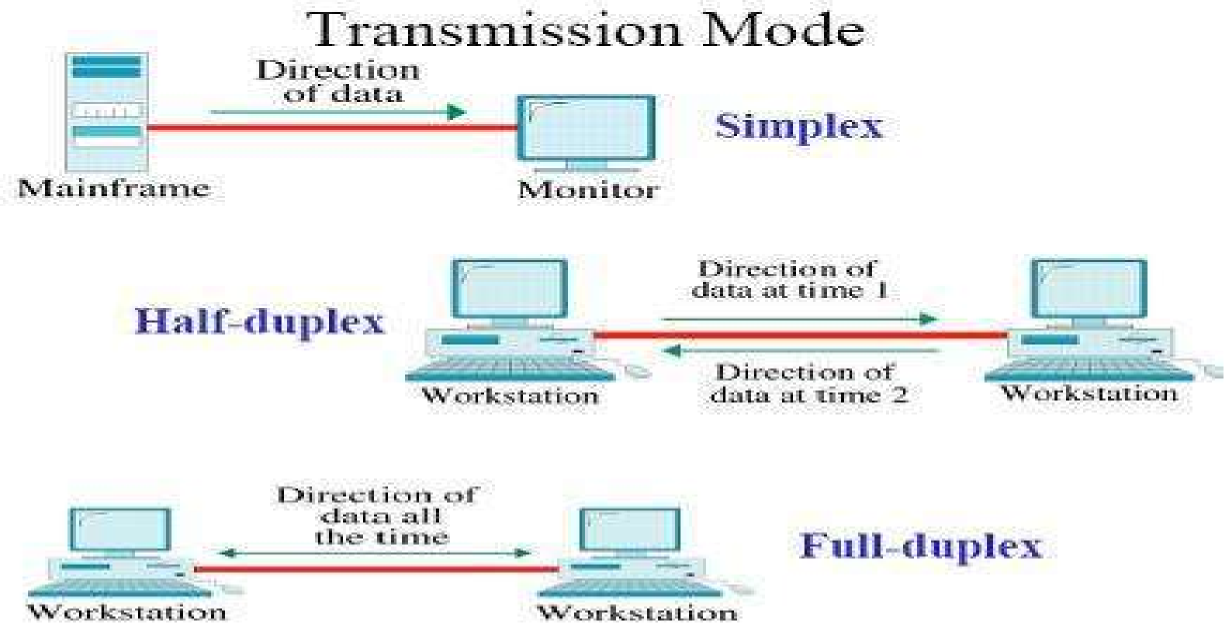
In full duplex transmission mode, the communication between sender and receiver can occur simultaneously. The sender and receiver can both transmit and receive at the same time. In this mode, signals can be sent in either direction. Full duplex transmission mode is like a two-way road, in which traffic can flow in both directions at the same time.



The sharing can be occur in two ways i.e. :

1) The link must contain 2 physical separate paths (one for sending and other for receiving)

2) The capacity of the channel is divided between signals travelling in opposite directions

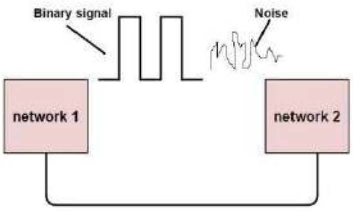


4.2 ERROR AND ERROR DETECTION :

ERROR :

The data can be corrupted during transmission (from source to receiver). It may be affected by external noise or some other physical imperfections. In this case, the input data is not same as the received output data. This mismatched data is called “Error”.

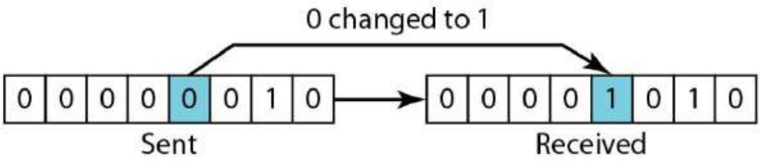
The data errors will cause loss of important / secured data. Even one bit of change in data may affect the whole system’s performance. Generally, the data transfer in digital systems will be in the form of ‘Bit – transfer’. In this case, the data error is likely to be changed in positions of 0 and 1



Types Of Errors
In a data sequence, if 1 is changed to zero or 0 is changed to 1, it is called “Bit error”. There are generally 3 types of errors occur in data transmission from transmitter to receiver. They are
- Single bit errors
- Multiple bit errors
- Burst errors

Single Bit Data Errors
The change in one bit in the whole data sequence , is called “Single bit error”. Occurrence of single bit error is very rare in serial communication system. Single bit errors are the least likely type error in serial transmission. Suppose, a

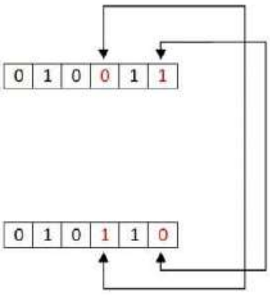
sender sends data of 1Mbps. It means each bit lasts on 1/10,00,000 seconds or 1µs. To occur a single bit error, the noise must have a duration of only 1µs.



This type of error occurs only in parallel communication system, as data is transferred bit wise in single line, there is chance that single line to be noisy.

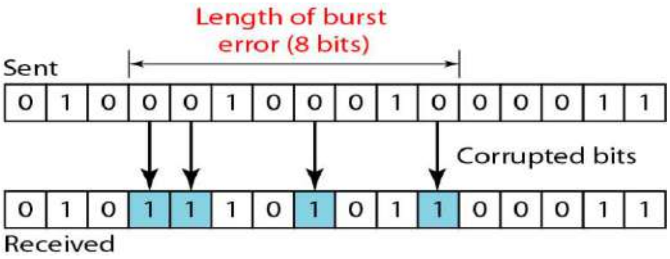
Multiple Bit Data Errors

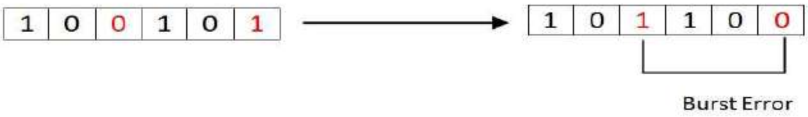
If there is change in two or more bits of data sequence of transmitter to receiver, it is called "Multiple bit error". This type of error occurs in both serial type and parallel type data communication networks.



Burst Errors

The term burst error means, 2 or more bits are changed from 0 to 1 or 1 to 0. The change of set of bits in data sequence is called "Burst error". The length of the burst is measured from the first corrupted bit to the last corrupted bit. This type of errors occurs in serial communication and they are difficult to solve.





The duration of noise is normally longer than the duration of a bit. It means, when noise affects data, it affects a set of bits. The number of bits affected depends on the data rate and duration of noise.

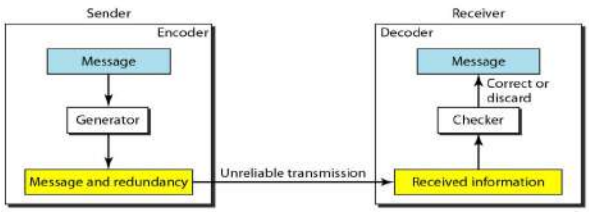
Example : If we sent data at 1Kbps, a noise of 1/00 sec can affect 10bits. If we sent data at 1Mbps, a noise can affect 10,000 bits.

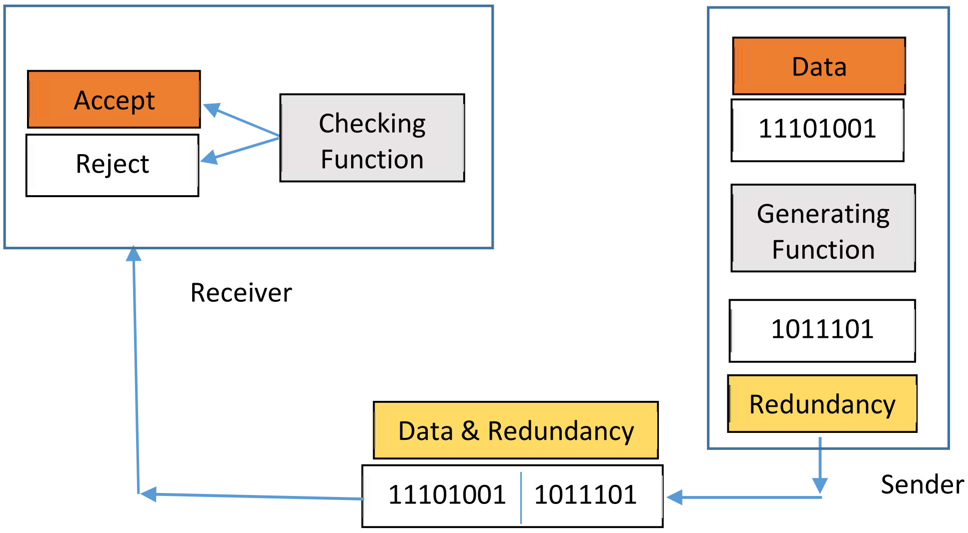
4.2 ERROR DETECTION :

Error detection is the process of detecting the errors that are present in the data transmitted from transmitter to receiver, in a communication system. One error detection method is to send every data unit twice. Then, the receiving device make bit-by-bit comparison. This process of re-transmission creates transmission time double and also comparison time also increase.

To avoid this method, a shorter group of bits may be appended to the end of each unit. This technique is called redundancy. We use some redundancy codes to detect these errors, by adding to the data while it is transmitted from source (transmitter)

In this method, extra bits are redundant to the information. These bits are discarded as soon as the accuracy of the transmission has been determined. Error detection uses the concept of redundancy, which means adding extra bits for detecting errors at the destination.





Types Of Redundancy Check :
There are four types redundancy check. These are :

1) VRC (Vertical Redundancy Check)
2) LRC (Longitudinal Redundancy Check)
3) CRC (Cyclical Redundancy Check)
4) Checksum

1) VRC (Vertical Redundancy Check)

Vertical Redundancy Check is also known as Parity Check.

The most common and least expensive mechanism for error detection if VRC. It is usually called Parity Check. In this technique, a redundant bit is called “Parity Bit” is appended to every data unit. This method includes even parity and odd parity. Even parity means the total number of 1s in data is to be even and odd parity means the total number of 1s in data is to be odd.

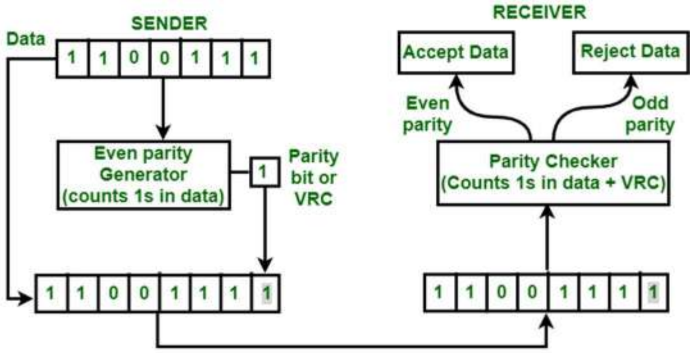
In even parity, the data unit is passed through the even parity generator. It counts the number of 1s in the data unit. If odd number of 1s, then it sets 1 in the parity bit to make the number of 1s as even. If the data unit having even number of 1s then it sets in the parity bit to maintain the number of 1s as even. When it reaches its destination, the receiver puts all bits through an even parity checking function. If it counts even number of 1s than there is no error. Otherwise there is some error.

EXAMPLE:
The data is : 01010110 (V)
The VRC check : 010101100

In odd parity, the data unit is passed through the odd parity generator. It counts the number of 1s in the data unit. If even number of 1s, then it sets 1 in the parity bit to make the number of 1s as odd. If the data unit having odd number of 1s then it sets in the parity bit to maintain the number of 1s as odd. When it reaches its destination, the receiver puts all bits through an odd parity checking function. If it counts odd number of 1s than there is no error. Otherwise there is some error.

EXAMPLE

The data is: 01010110 (V)
The VRC check: 01010111



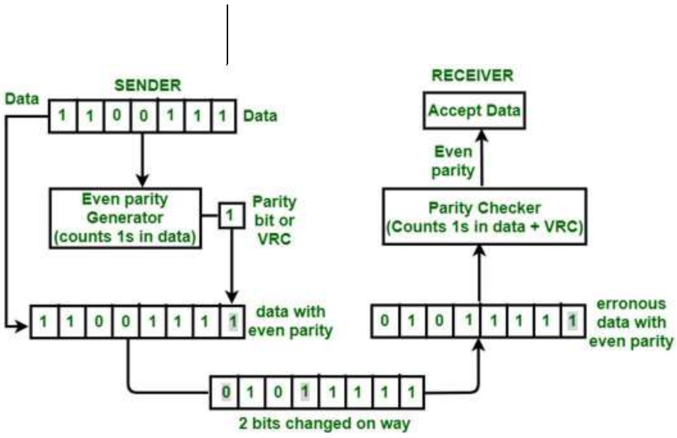
Advantages :

- VRC can detect all single bit error.
- It can also detect burst errors but only in those cases where number of bits changed is odd, i.e. 1, 3, 5, 7, …….etc.

Disadvantages :

The major disadvantage of using this method for error detection is that it is not able to detect burst error if the number of bits changed is even, i.e. 2, 4, 6, 8, …….etc.

Example –
If the original data is 1100111. After adding VRC, data unit that will be transmitted is 11001111. Suppose on the way 2 bits are 01011111. When this data will reach the destination, parity checker will count number of 1s in data and that comes out to be even i.e. 8. So, in this case, parity is not changed, it is still even. Destination will assume that there is no error in data even though data is erroneous.

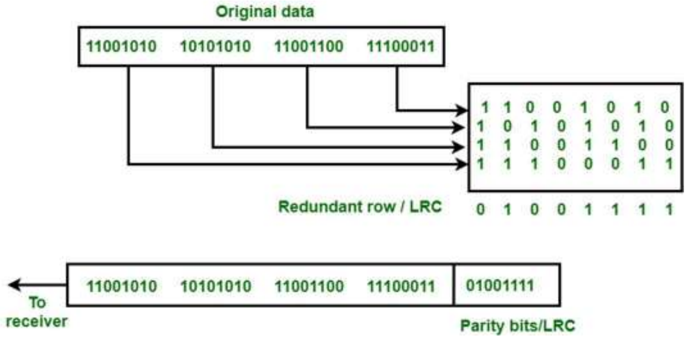


Erroneous data accepted by receiver with number of changed bits

2) LRC (Longitudinal Redundancy Check)

Longitudinal Redundancy Check (LRC) is also known as 2-D parity check. In this method, data is organised into tables of rows and columns. A block of bit is divided into table or matrix of rows and columns. In order to detect an error, a redundant bit is added to the whole block and this block is transmitted to receiver. The receiver uses this redundant row to detect error. After checking the data for errors, receiver accepts the data and discards the redundant row of bits.

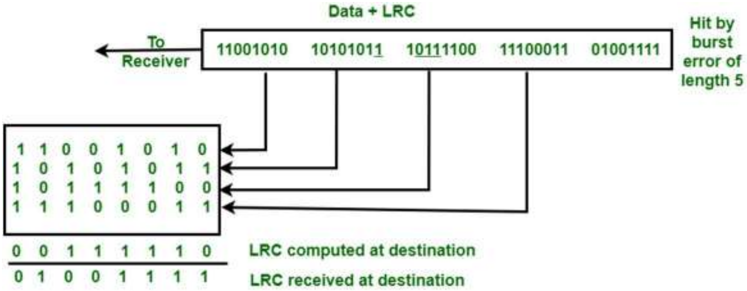
Example :
If a block of 32 bits is to be transmitted, it is divided into matrix of four rows and eight columns which as shown in the following figure :



In this matrix of bits, a parity bit (odd or even) is calculated for each column. It means 32 bits data plus 8 redundant bits are transmitted to receiver. Whenever data reaches at the destination, receiver uses LRC to detect error in data.

Advantage :
LRC is used to detect burst errors.

Example : Suppose 32 bit data plus LRC that was being transmitted is hit by a burst error of length 5 and some bits are corrupted as shown in the following figure :



The LRC received by the destination does not match with newly corrupted LRC. The destination comes to know that the data is erroneous, so it discards the data.

Disadvantage :
The main problem with LRC is that, it is not able to detect error if two bits in a data unit are damaged and two bits in exactly the same position in other data unit are also damaged.
Example : If data 110011 010101 is changed to 010010110100.

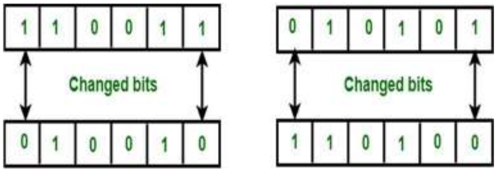


Figure : *Two bits at same bit position damaged in 2 data units*

In this example 1st and 6th bit in one data unit is changed . Also the 1st and 6th bit in second unit is changed.

3) CRC (Cyclical Redundancy Check)

The third and most powerful of the redundancy checking techniques is the CRC (Cyclical Redundancy Check). It is based on the binary division. In CRC, a sequence of redundant bits called CRC/ the CRC reminder, are added to the end of a data unit. So that, the resulting data unit becomes exactly divisible by a second, predetermined binary number.

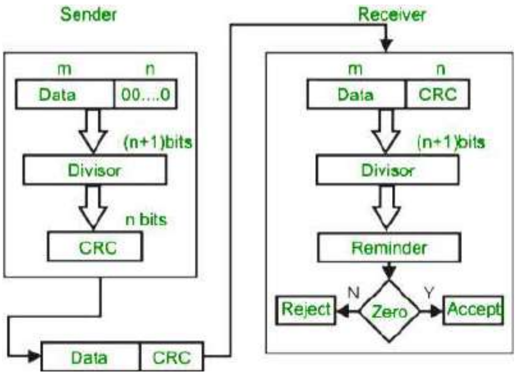
CRC is a method of detecting accidental changes/errors in the communication channel.
CRC uses **Generator Polynomial** which is available on both sender and receiver side. An example generator polynomial is of the form like $x^3 + x + 1$. This generator polynomial represents key 1011. Another example is $x^2 + 1$ that represents key 101.

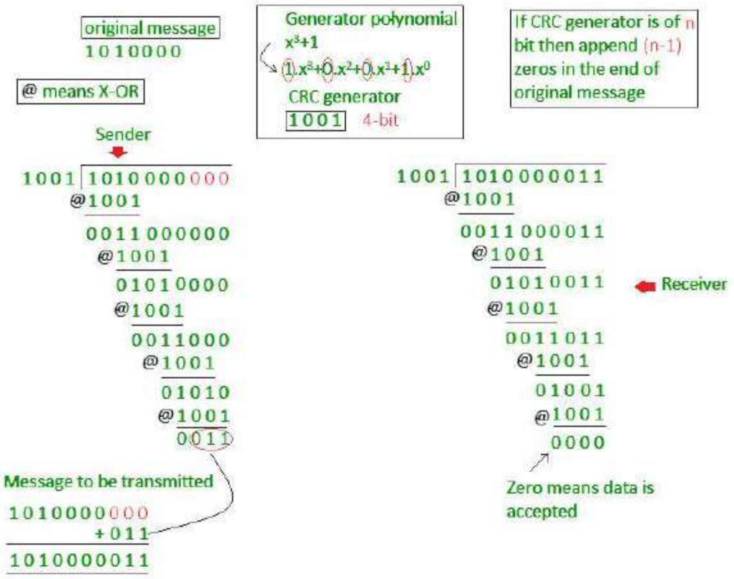
At the destination, the incoming data unit is divided by the same number. If at this step there is no remainder, the data unit is assumed to be correct and is therefore accepted. A remainder indicates that the data unit has been damaged in transit and therefore must be rejected.

The redundancy bits used by CRC are derived by dividing the data unit by a predetermined divisor. The reminder is the CRC. To be a valid, CRC must have two(02) qualities :

* It must have exactly one less bit than the divisor

* Appending it to the end of the data string to make the resulting bit sequence exactly divisible by the divisor.





CRC Generator :

A CRC generator uses moduloe-2 division. In the 1st step, the 4-bitdvisor is subtracted from First 4bits of the dividend. Each bit of the divisor is subtracted from the corresponding bit of the dividend without disturbing the next higher bit.

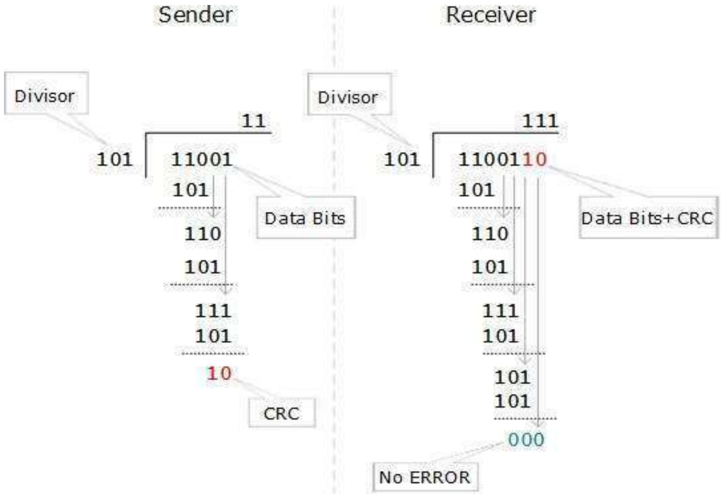
CRC Checker :

A CRC checker functions exactly like the generator. After receiving the data added with the CRC, it does the same modulo-2 division. If the reminder is all 0s, the CRC is dropped and the data is accepted. Otherwise, the received bit stream is discarded

Modulo 2 Division:

The process of modulo-2 binary division is the same as the familiar division process we use for decimal numbers. Just that instead of subtraction, we use XOR here.

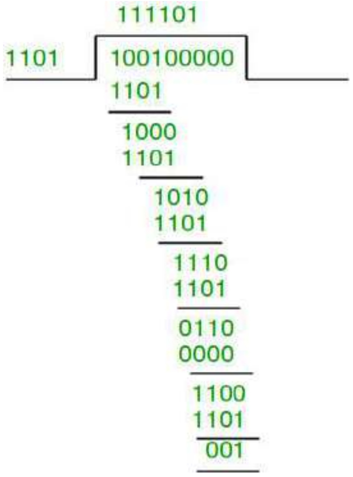
- In each step, a copy of the divisor (or data) is XORed with the k bits of the dividend (or key).
- The result of the XOR operation (remainder) is (n-1) bits, which is used for the next step after 1 extra bit is pulled down to make it n bits long.
- When there are no bits left to pull down, we have a result. The (n-1)-bit remainder which is appended at the sender side.



Example 1 (No error in transmission):

- Data word to be sent - 100100

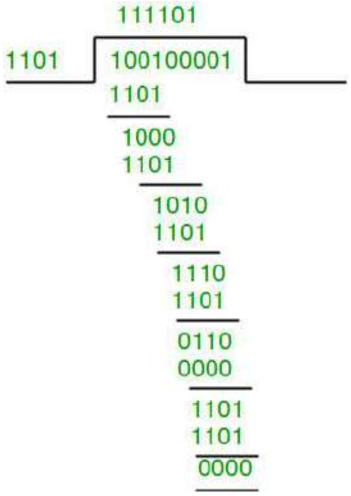
- Key - 1101 [Or generator polynomial x3 + x2 + 1]



- Therefore, the remainder is 001 and hence the encoded
- data sent is 100100001.

Receiver Side:

- Code word received at the receiver side 100100001



Therefore, the remainder is all zeros. Hence, the data received has no error.

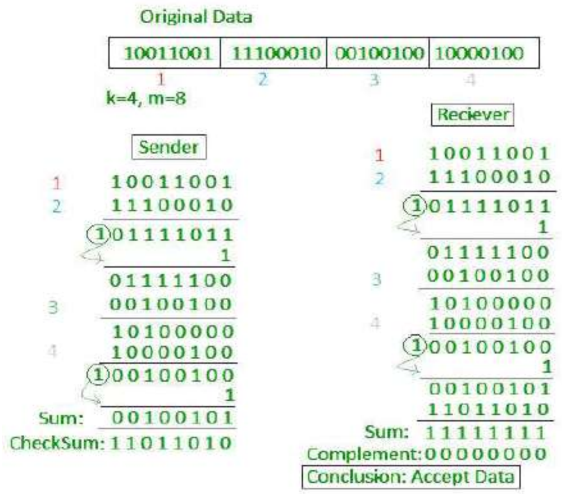
4) Checksum

The error detection method used by the higher layer protocol is called CHECKSUM. Like VrC, LrC and CRC, it is based on the concept of redundancy.

- In checksum error detection scheme, the data is divided into k segments each of m bits.
- In the sender's end the segments are added using 1's complement arithmetic to get the sum. The sum is complemented to get the checksum.
- The checksum segment is sent along with the data segments.
- At the receiver's end, all received segments are added using 1's complement arithmetic to get the sum. The sum is complemented.
- If the result is zero, the received data is accepted; otherwise discarded.

A checksum is a small-sized block of data derived from another block of digital data for the purpose of detecting errors. Checksums are often used to verify data integrity but are not relied upon to verify data authenticity.

This is a block code method where a checksum is created based on the data values using some algorithm and appended to the data. When the receiver gets this data, a new checksum is calculated and compared with the existing checksum. A non-match indicates an error.



Checksum Generator :

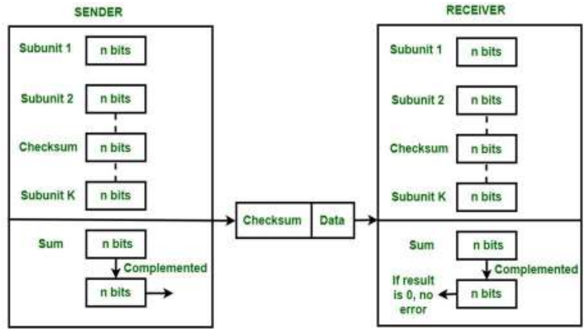
The checksum generator subdivide the data unit into equal segments of n-bits (usually 16). These segments are added together using 1’s complement arithmetic.

Steps followed by Sender :

- The unit is divided into “k” section, each of n-bits
- All sections are added together using 1’s complement to get the sum
- The sum is complemented and becomes the checksum
- The checksum is sent with the data

Steps followed by Receiver :

- The unit is divided into “k” section, each of n-bits
- All sections are added together using 1’s complement to get the sum
- The sum is complemented
- If the result is 0(zero), the data is accepted otherwise rejected.



Example :

Data Accepted

A block of 16 bits is to be sent using a checksum of 8bts.

169 and 57

The ASCII values are : 10101001 and 00111001

The numbers are added using 1’s complement arithmetic :

10101001

00111001

Sum : 11100010

Checksum: 00011101

The pattern sent is : 10101001 00111001 00011101 (checksum)

At the receiver side, the receiver then add the 3 sections together. After complement, if result is 0(zero), then it is accepted.

10101001

00111001

00011101

Sum : 11111111

Checksum: 00000000 (Accepted)

Data Rejected

A block of 16 bits is to be sent using a checksum of 8bts.

169 and 57

The ASCII values are : 10101001 and 00111001

The numbers are added using 1’s complement arithmetic :

10101001
00111001
Sum : 11100010
Checksum: 00011101

The pattern sent is : 10101001 00111001 00011101 (checksum)

Data changed (during transmission: 10101**11**1 **11**111001 00011101 (checksum)

At the receiver side, the receiver then add the 3 sections together. After complement, if result is 0(zero), then it is accepted.

10101**11**1
11111001
00011101
Result : 11000101
Carry 1
Sum : 11000110
Checksum: 00111001 (Rejected)

4.3 LINE CONFIGURATION :

Line configuration refers to the way, two or more communication devices attached to a link. Line configuration is also referred to as connection. A link is the physical communication pathway that transfers data from one device to another. For communication to occur, two devices must be connected in same way to the same link at the same time.

There are 2 possible line configurations. These are :

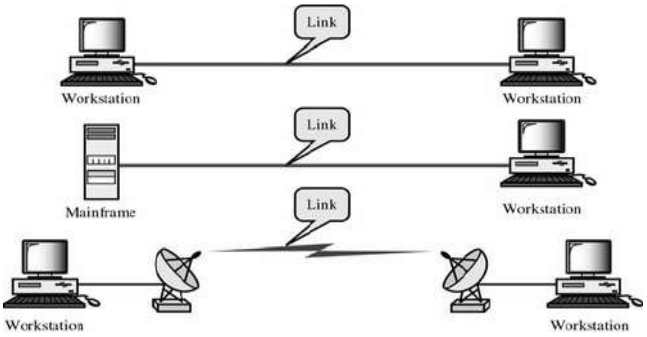
i) Point-to-Point
ii) Multipoint

Point-to-Point

A Point to Point Line Configuration provide dedicated link between two devices. The entire capacity of the channel is reserved for transmission between those two devices.

Most point-to-point line configurations use an actual length of wire or cable to connect the two ends, but other options, such as microwave or satellite links, TV remote are also possible.

Point to point network topology is considered to be one of the easiest and most conventional network topologies. It is also the simplest to establish and understand. To visualize, one can consider point to point network topology as two phones connected end to end for a two way communication



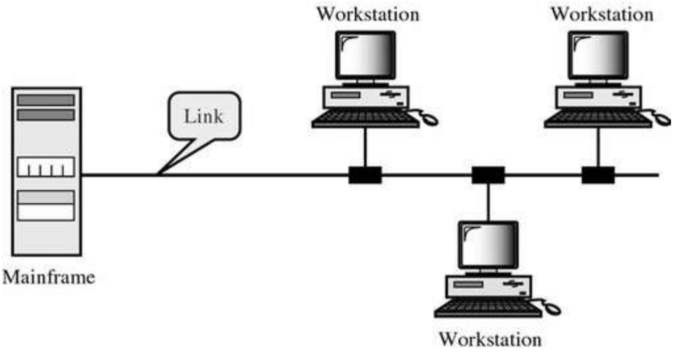
Multipoint Configuration

Multipoint Configuration also known as Multidrop line configuration. In this configuration, one or more than two   devices share a single link capacity of the channel.

More than two devices share the Link that is the capacity of the channel is shared now. With shared capacity, there can be two possibilities in a Multipoint Line Config:

- **Spatial Sharing**: If several devices can share the link simultaneously, its called Spatially shared line configuration

- **Temporal (Time) Sharing**: If users must take turns using the link , then its called Temporally shared or Time Shared Line Configuration



4.4 FLOW CONTROL :

Flow control is a technique that allows two stations working at different speeds to communicate with each other. It is a set of measures taken to regulate the amount of data that a sender sends so that a fast sender does not overwhelm a slow receiver. In data link layer, flow control restricts the number of frames the sender can send before it waits for an acknowledgment from the receiver.

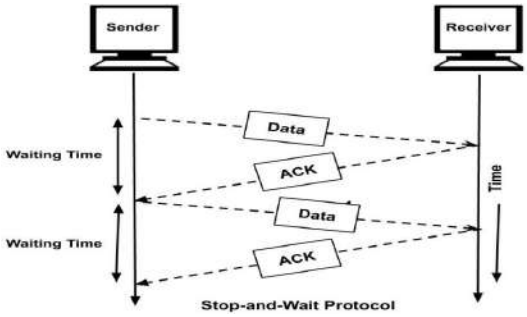
- It is a set of procedures that tells the sender how much data it can transmit before the data overwhelms the receiver.
- The receiving device has limited speed and limited memory to store the data. Therefore, the receiving device must be able to inform the sending device to stop the transmission temporarily before the limits are reached.
- It requires a buffer, a block of memory for storing the information until they are processed.

Two methods have been developed to control the flow of data:

- Stop-and-wait
- Sliding window

Stop-and-wait

- In the Stop-and-wait method, the sender waits for an acknowledgement after every frame it sends.
- When acknowledgement is received, then only next frame is sent. The process of alternately sending and waiting of a frame continues until the sender transmits the EOT (End of transmission) frame.



Stop-and-Wait Protocol

In this method, the sender sends one frame and waits for an ACK before sending the next frame. Each frame must travel all the way to the receiver. An ACK must travel all the way back before the next frame can be sent.

Advantage of Stop-and-wait

The Stop-and-wait method is simple as each frame is checked and acknowledged before the next frame is sent.

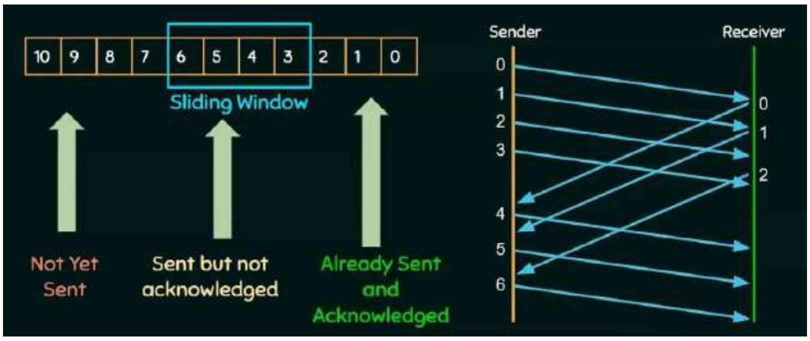
Disadvantage of Stop-and-wait

Stop-and-wait technique is inefficient to use as each frame must travel across all the way to the receiver, and an acknowledgement travels all the way before the next frame is sent. Each frame sent and received uses the entire time needed to traverse the link.

Sliding Window

- The Sliding Window is a method of flow control in which a sender can transmit the several frames before getting an acknowledgement.

- In Sliding Window Control, multiple frames can be sent one after the another due to which capacity of the communication channel can be utilized efficiently.
- A single ACK acknowledge multiple frames.
- Sliding Window refers to imaginary boxes at both the sender and receiver end.
- The window can hold the frames at either end, and it provides the upper limit on the number of frames that can be transmitted before the acknowledgement.
- Frames can be acknowledged even when the window is not completely filled.
- The window has a specific size in which they are numbered as modulo-n means that they are numbered from 0 to n-1. For example, if n = 8, the frames are numbered from 0,1,2,3,4,5,6,7,0,1,2,3,4,5,6,7,0,1........
- The size of the window is represented as n-1. Therefore, maximum n-1 frames can be sent before acknowledgement.
- When the receiver sends the ACK, it includes the number of the next frame that it wants to receive. For example, to acknowledge the string of frames ending with frame number 4, the receiver will send the ACK containing the number 5. When the sender sees the ACK with the number 5, it got to know that the frames from 0 through 4 have been received.



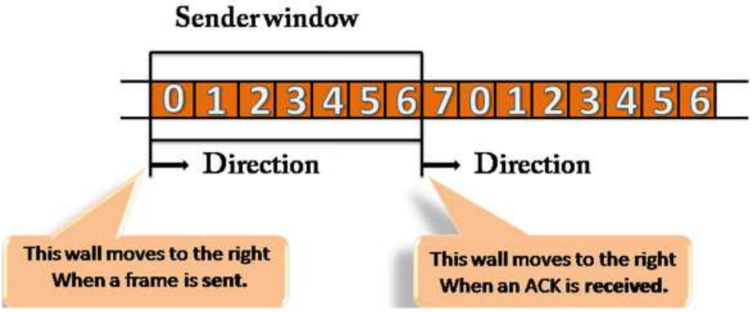
Sender Window

- At the beginning of a transmission, the sender window contains n-1 frames.
- As the frames are sent out, the left boundary moves inward.
- It shrinks the size of the window.

  For example, if the size of the window is w if three frames are sent out, then the number of frames left out in the sender window is w-3.

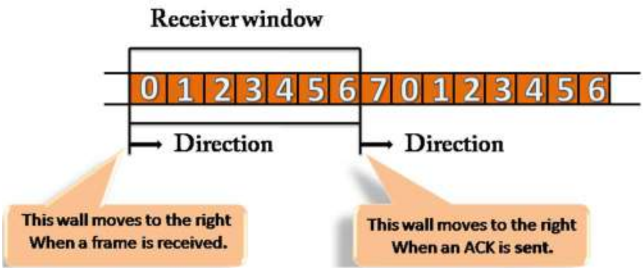
- Once the ACK has arrived, then the sender window expands to the number which will be equal to the number of frames acknowledged by ACK.

  For example, the size of the window is 7, and if frames 0 through 4 have been sent out and no acknowledgement has arrived, then the sender window contains only two frames, i.e., 5 and 6. Now, if ACK has arrived with a number 4 which means that 0 through 3 frames have arrived undamaged and the sender window is expanded to include the next four frames. Therefore, the sender window contains six frames (5,6,7,0,1,2).


Senderwindow
0 1 2 3 4 5 6 7 0 1 2 3 4 5 6
Direction
Direction
This wall moves to the right When a frame is sent.
This wall moves to the right When an ACK is received.

Receiver Window

- ○ At the beginning of transmission, the receiver window does not contain n frames, but it contains n-1 spaces for frames.
- ○ When the new frame arrives, the size of the window shrinks.
- ○ The receiver window does not represent the number of frames received, but it represents the number of frames that can be received before an ACK is sent. For example, the size of the window is w, if three frames are received then the number of spaces available in the window is (w-3).
- ○ Once the acknowledgement is sent, the receiver window expands by the number equal to the number of frames acknowledged.
- ○ Suppose the size of the window is 7 means that the receiver window contains seven spaces for seven frames. If the one frame is received, then the receiver window shrinks and moving the boundary from 0 to 1. In this way, window shrinks one by one, so window now contains the six spaces. If frames from 0 through 4 have sent, then the window contains two spaces before an acknowledgement is sent.


Receiverwindow
0 1 2 3 4 5 6 7 0 1 2 3 4 5 6
Direction
Direction
This wall moves to the right When a frame is received.
This wall moves to the right When an ACK is sent.

4.5 ERROR CORRECTION :

Error Control is a technique of error detection and retransmission. It can be handled in 2 ways.

1. When an error is discovered, the receiver can be the sender to retransmit the entire data unit.
2. A receiver can use an error correction code which automatically corrects the error.

Error correction codes are more sophisticated than error detection code. It requires more redundancy bits. The number of bits required to correct a multi-bit or burst error is so high. For this reason, most error correction is limited to 1-bit, 2-bit or 3-bit errors.

Single bit/1-bit Error Control :

A single bit error can be detected by the addition of a redundancy bit (parity bit) to the data unit. A bit has two states (0 & 1). An error occurs when the receiver reads 1-bit as 0 and 0-bit as 1.

To correct the error, the receiver simply revers the value of the altered bit. To do so, it must know which bit is in error.

Example :

To correct a single bit error in an ASCII character, the error correction code must determine which of the 7bits has changed. For this, it has to distinguish between 8 different states i.e. no error, error in 1st position, error in 2nd position, and so on upto 7th positon. To do so, it requires enough redundancy bits to show all 8 states.

At first, it will look a 3 bit redundancy code. Because, 3-bits can show 8 different states i.e. 000 to 111. It indicates the location of 8 different possibilities.

But in case of redundancy bit error, (7bit of data i.e. ASCII + 3 bits of redundancy bits), 3 bits can identify only 8 possibilities

Redundancy Bit :

Redundant bits are extra binary bits that are generated and added to the information-carrying bits of data transfer to ensure that no bits were lost during the data transfer. To calculate the number of redundancy bits (r) required to correct a given number of data bits (m), we must find out a relationship between them.

If total number of bits in a transmittable unit is m+r, then r must be able to indicate at least m+r+1 different states. Here, m+r states indicates the location of an error in each m+r position one 1 indicates no error.

Therefore, 2^r>=m+r+1. The value of r can be determined by plugging in the value of m.

Example :

If the value of m is 7, the smallest value of r to satisfy the equation is 4 i.e $2^4 >= 7 + 4 + 1$

Parity bits –
A parity bit is a bit appended to a data of binary bits to ensure that the total number of 1's in the data is even or odd. Parity bits are used for error detection. There are two types of parity bits:

1. ***Even parity bit:***
   In the case of even parity, for a given set of bits, the number of 1's are counted. If that count is odd, the parity bit value is set to 1, making the total count of occurrences of 1's an even number. If the total number of 1's in a given set of bits is already even, the parity bit's value is 0.
2. ***Odd Parity bit –***
   In the case of odd parity, for a given set of bits, the number of 1's are counted. If that count is even, the parity bit value is set to 1, making the total count of occurrences of 1's an odd number. If the total number of 1's in a given set of bits is already odd, the parity bit's value is 0.

Hamming Code :

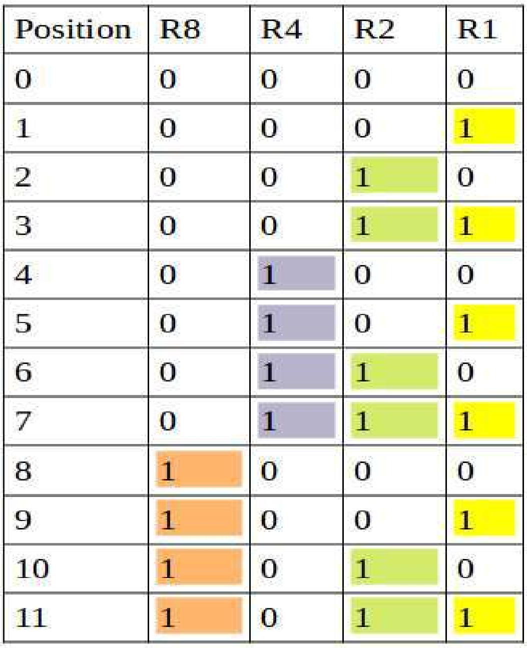
This technique was developed by R. W. Hamming. The Hamming Code can be applied to data units of any length and uses the relationship between data and redundancy bits.

Hamming code is a set of error-correction codes that can be used to detect and correct the errors that can occur when the data is moved or stored from the sender to the receiver. It is technique developed by R.W. Hamming for error correction.

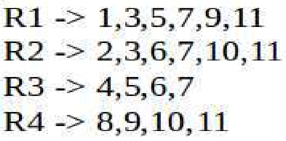
General Algorithm of Hamming code –

The Hamming Code is simply the use of extra parity bits to allow the identification of an error.

1. Write the bit positions starting from 1 in binary form (1, 10, 11, 100, etc).
   0001, 0010, 0011, 0100, 0101, 0110, 0111,, 1000, 1001, 1010, 1011, 1100, 1101, 1110, 1111
2. All the bit positions that are a power of 2 are marked as parity bits (1, 2, 4, 8, etc).
3. All the other bit positions are marked as data bits.
4. Each data bit is included in a unique set of parity bits, as determined its bit position in binary form.
   a. **Parity bit 1** covers all the bits positions whose binary representation includes a 1 in the least significant position (1, 3, 5, 7, 9, 11, etc).
   b. **Parity bit 2** covers all the bits positions whose binary representation includes a 1 in the second position from the least significant bit (2, 3, 6, 7, 10, 11, etc).
   c. **Parity bit 4** covers all the bits positions whose binary representation includes a 1 in the third position from the least significant bit (4–7, 12–15, 20–23, etc).
   d. **Parity bit 8** covers all the bits positions whose binary representation includes a 1 in the fourth position from the least significant bit bits (8–15, 24–31, 40–47, etc).
   e. In general, each parity bit covers all bits where the bitwise AND of the parity position and the bit position is non-zero.
5. Since we check for even parity set a parity bit to 1 if the total number of ones in the positions it checks is odd.
6. Set a parity bit to 0 if the total number of ones in the positions it checks is even.



Position	R8	R4	R2	R1
0	0	0	0	0
1	0	0	0	1
2	0	0	1	0
3	0	0	1	1
4	0	1	0	0
5	0	1	0	1
6	0	1	1	0
7	0	1	1	1
8	1	0	0	0
9	1	0	0	1
10	1	0	1	0
11	1	0	1	1



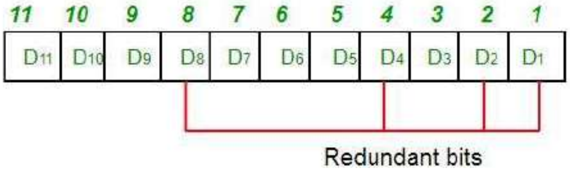
R1 -> 1,3,5,7,9,11
R2 -> 2,3,6,7,10,11
R3 -> 4,5,6,7
R4 -> 8,9,10,11

Determining the position of redundant bits –

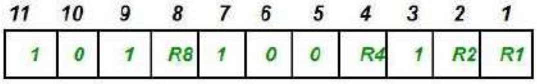
These redundancy bits are placed at the positions which correspond to the power of 2.

As in the above example:

1. The number of data bits = 7
2. The number of redundant bits = 4
3. The total number of bits = 11
4. The redundant bits are placed at positions corresponding to power of 2- 1, 2, 4, and 8

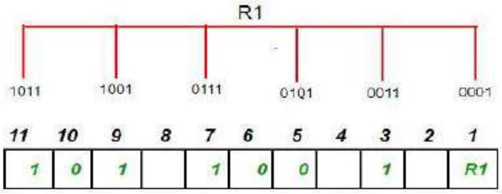


Suppose the data to be transmitted is 1011001, the bits will be placed as follows:



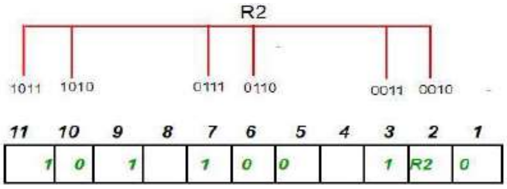
Determining the Parity bits –

1. **R1** bit is calculated using parity check at all the bits positions whose binary representation includes a 1 in the least significant position.
   R1: bits 1, 3, 5, 7, 9, 11

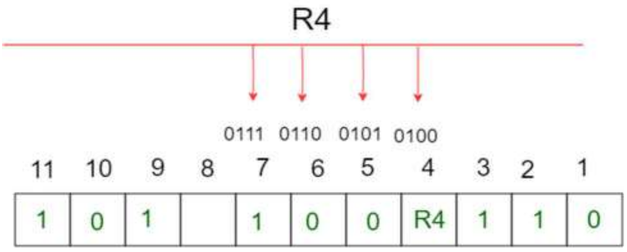
   To find the redundant bit R1, we check for even parity. Since the total number of 1's in all the bit positions corresponding to R1 is an even number the value of R1 (parity bit's value) = 0

2. **R2** bit is calculated using parity check at all the bits positions whose binary representation includes a 1 in the second position from the least significant bit.
   R2: bits 2,3,6,7,10,11

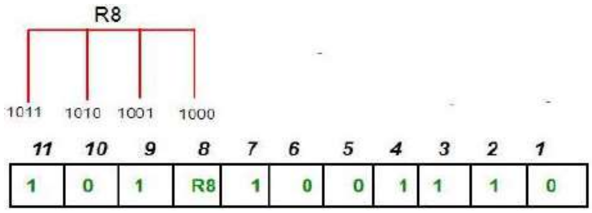
   To find the redundant bit R2, we check for even parity. Since the total number of 1's in all the bit positions corresponding to R2 is odd the value of R2(parity bit's value)=1

3. **R4** bit is calculated using parity check at all the bits positions whose binary representation includes a 1 in the third position from the least significant bit.
   R4: bits 4, 5, 6, 7



To find the redundant bit R4, we check for even parity. Since the total number of 1's in all the bit positions corresponding to R4 is odd the value of R4(parity bit's value) = 1

4. **R8** bit is calculated using parity check at all the bits positions whose binary representation includes a 1 in the fourth position from the least significant bit.
R8: bit 8,9,10,11

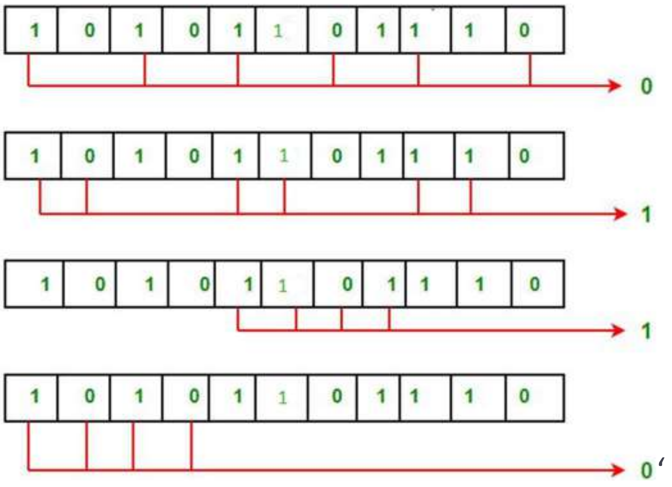


To find the redundant bit R8, we check for even parity. Since the total number of 1's in all the bit positions corresponding to R8 is an even number the value of R8(parity bit's value)=0.

Thus, the data transferred is:

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

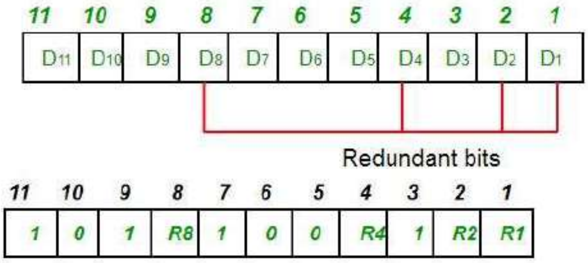
Error detection and correction –
Suppose in the above example the 6th bit is changed from 0 to 1 during data transmission, then it gives new parity values in the binary number:



The bits give the binary number as 0110 whose decimal representation is 6. Thus, the bit 6 contains an error. To correct the error the 6th bit is changed from 1 to 0.

Example :

A 7-bit ASCII core requires 4 redundancy bits that can be added to the end of the data unit or interspread wth the original data bits. Positions of redundancy bits in Hamming code :



Determining the position of redundant bits –

These redundancy bits are placed at the positions which correspond to the power of 2.

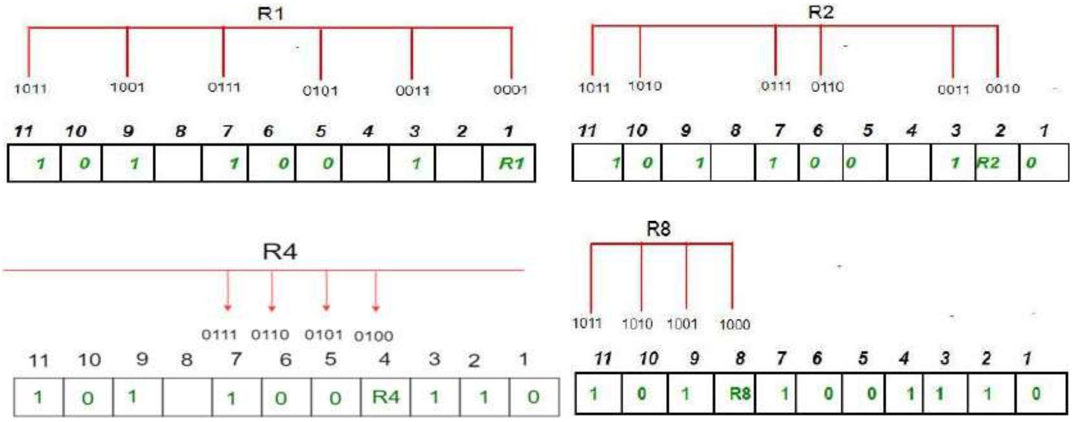
As in the above example:

1. The number of data bits = 7
2. The number of redundant bits = 4
3. The total number of bits = 11
4. The redundant bits are placed at positions corresponding to power of 2- 1, 2, 4, and 8

In Hamming code, each r bit s the VRC bit for one combination of data bits. R1 is the VRC bit for one combination, R2 is the VRC for another combination and so on. The combination used to calculate each of the 4 R values for a seven (7) bit data sequence as :

R1 -> 1,3,5,7,9,11
R2 -> 2,3,6,7,10,11
R3 -> 4,5,6,7
R4 -> 8,9,10,11

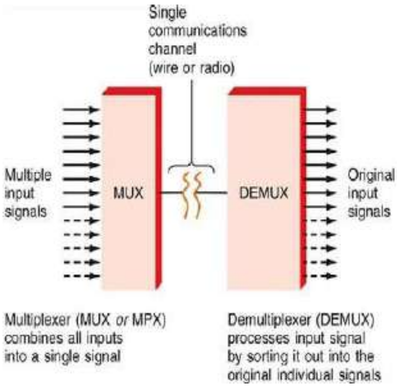
This strategy is based on the binary representation of each bit position. The R bit is calculated using all bit positions whose binary representation includes 1 in the right most position. The R2 bit is calculated using all bit positions with a 1 in the 2nd position and so on.



4.6 MULTIPLEXING :

- Multiplexing is a technique by which different multiple signal streams can be transmitted simultaneously over a shared link or single data 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.


Single communications channel (wire or radio)
Multiple input signals
MUX
DEMUX
Original input signals
Multiplexer (MUX or MPX) combines all inputs into a single signal
Demultiplexer (DEMUX) processes input signal by sorting it out into the original individual signals

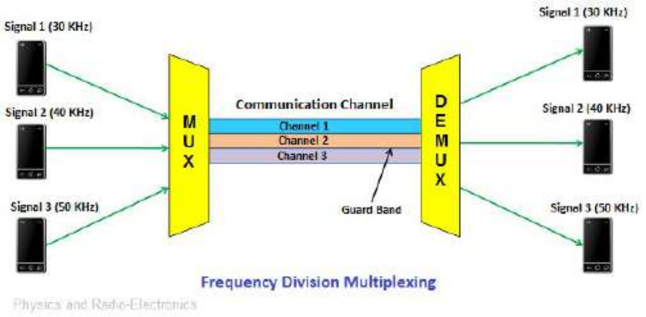
- When more than one senders tries to send over single medium, a device called Multiplexer divides the physical channel and allocates one to each. On the other end of communication, a De-multiplexer receives data from a single medium and identifies each and send to different receivers.
- Transmitting two or more signals simultaneously can be accomplished by running multiple cables or setting up one transmitter receiver pair for each channel , but this is an expensive approach.
- A single cable or radio link can handle multiple signals simultaneously using a technique known as multiplexing. Multiplexing permits hundreds or even thousands of signals to be combined and transmitted over a single medium.
- A device called a multiplexer (often shortened to "mux") combines the input signals into one signal. When the multiplexed signal needs to be separated into its component signal s (for example, when your email is to be delivered to its destination), a device called a d emultiplexer (or "demux") is used.
- Multiplexing was originally developed in the 1800s for telegraphy. Today, multiplexing is widely used in many telecommunications applications, including telephony, Internet com munications, digital broadcasting and wireless telephony.

Frequency Division Multiplexing

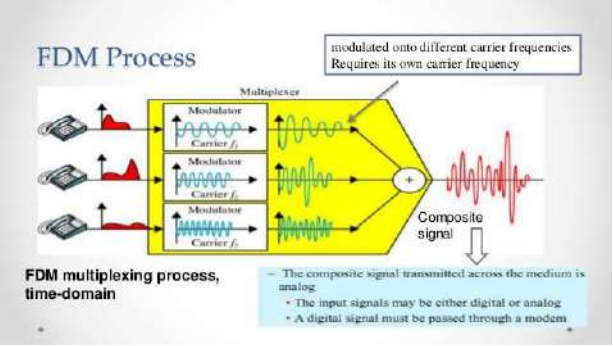
FDM is an analog technique. It can be applied when the bandwidth of a link is greater than the combined bandwidth of the signals to be transmitted. In FDS, signals generated by each sending devices, modulate different carrier frequencies. These modulated signals are then combined into a single composite signal that can be transmitted by a link.

Career frequencies are separated by enough bandwidth to accommodate the modulate signal. These bandwidth ranges are channels through which the various signals can travel. Channels must be separated by strips of unused bandwidth (guard bands) to prevent signals from overlapping.

Carrier frequencies must not interfere with the original data frequencies.

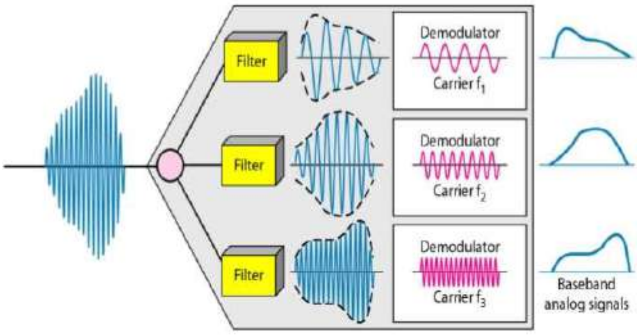


FDM Process :



FDM multiplexing process, time-domain

– The composite signal transmitted across the medium is analog
• The input signals may be either digital or analog
• A digital signal must be passed through a modem

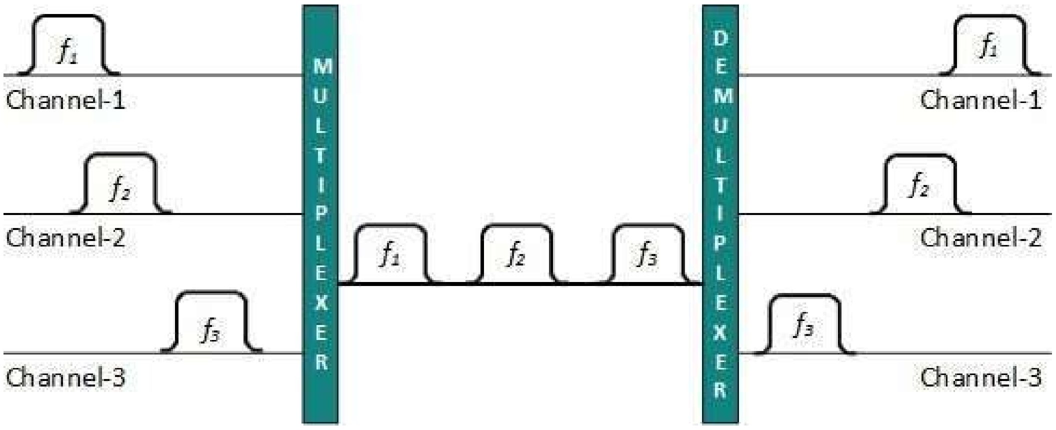


Time Domain FDM Multiplexing

In the above figures, FDM is an analog process. Each device generates a signal of a similar frequency range. Inside the multiplexer, these signals are modulated into different carrier frequencies (f1, f2, f3) which then combined into a single composite signal and send over a media link. The demultiplexer uses a series of filters to decompose the multiplexed signals into its constituent component signals.

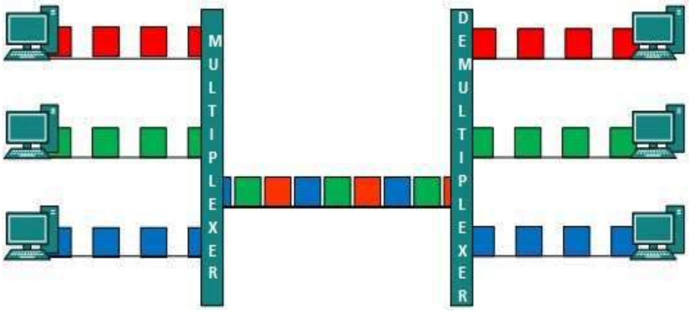
Then the individual signals are passed to a demodulator that separates them from their carriers and passes them to the waiting receivers.

- FDM is an analog technology.
- When the carrier is frequency, FDM is used.
- 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 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 aswell.
- 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 at one side channel A is transmitting its frame, on the other end De-multiplexer providing media to channel A.
- As soon as its channel A's time slot expires this side switches to channel B.
- On the other end De-multiplexer behaves in a synchronized manner and provides media to channel B. Signals from different channels travels the path in interleaved manner.
- Multiple signals can be transmitted if each signal is allowed to be transmitted for a definite amount of time. These time slots are so small that all transmissions appear to be in parallel.
- Further, on each wavelength Time division multiplexing can be incorporated to accommodate more data signals

TDM can be implemented in two (02) ways i.e.

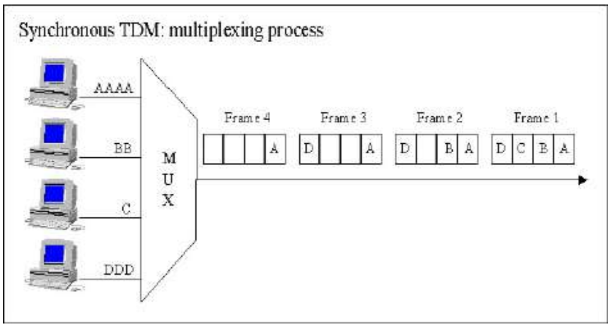
1) Synchronous TDM
2) Asynchronous TDM (Statistical TDM)

1. **Synchronous TDM:**

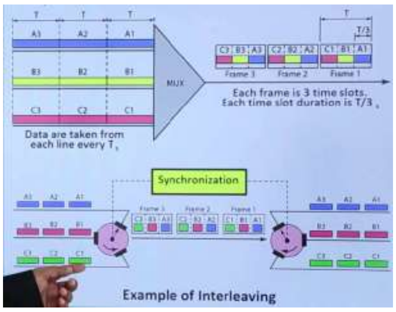
Here, synchronous means, the multiplexer allocates exactly the same time slot to each device at all times whether a device has anything to transmit or not.

Example :

Time slot “A” is assigned to device A along. It can not be used any other devices. Each time, its allocated time slot comes, A device has the opportunity to send a portion of its data. If A device is unable to transmit or does not have data to sent, its time slot remains empty.



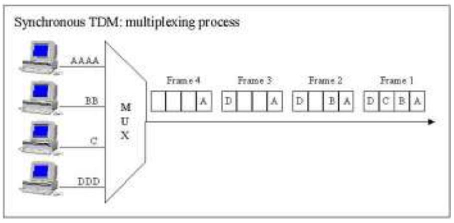
Time slots are pre assigned and are fixed. Each source is given its time slot at every turn due to it. This turn may be once per cycle, or several turns per cycle, if it has a high data transfer rate, or may be once in a no. of cycles if it is slow. This slot is given even if the source is not ready with data. So this slot is transmitted empty.
in the same manner as FDM but uses light as signals.



Frames :

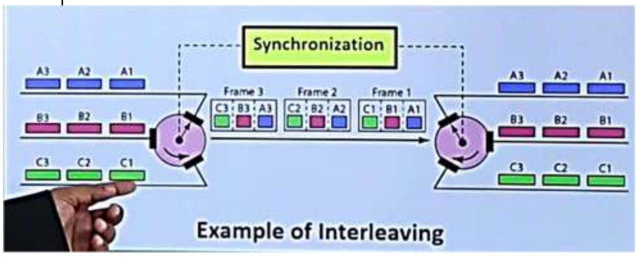
Time slots are grouped into frames. A frame consists of complete cycle of time slots. It includes one or more slots dedicated to each sending devices.

In a system with “n” input lines, each frame has at least “n” slots with each slot allocated to carrying data from a specific input line. It all the input devices sharing a link and transmitting at the same data rate, then each device has one time slot per frame.



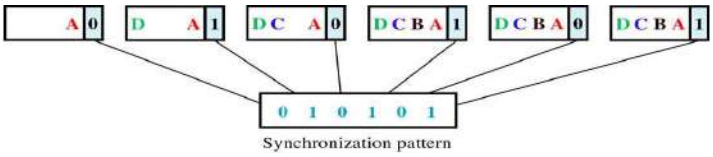
Inter leaving :

Synchronous TDM can be compared to a very fast rotating switch. As the switch opens in from of a device, that device has the opportunity to send a specified amount of data into the path. The switch moves from one device to another at a constant rate and in a fixed order. This process is called interleaving.



Frame Bits :

In synchronous TDM, the time slot order does not vary from frame to frame. Very little more/extra info^n needs to be included in each frame. Various factors cause timing inconsistencies. For this reasons, one or more synchronous bits are added to the beginning of each frame. These bits are called frame bits.



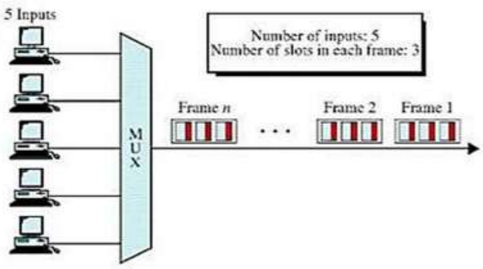
2. Asynchronous TDM:

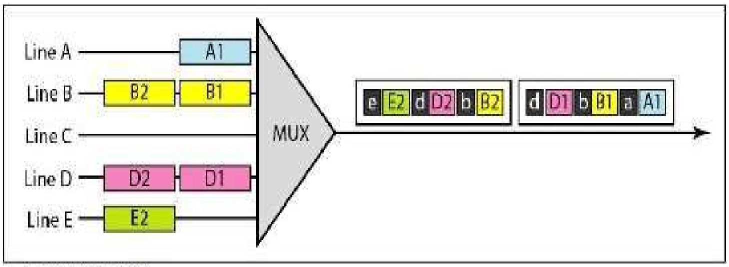
Asynchronous TDM is known as "Statistical TDM". Asynchronous means, flexible. In this method, slots are not fixed. They are allotted dynamically depending on speed of sources, and whether they are ready for transmission.

It allows a number of lower speed unit lines to be multiplexed to a single higher speed line. Total speed of the input lines can be greater than the capacity of the path.

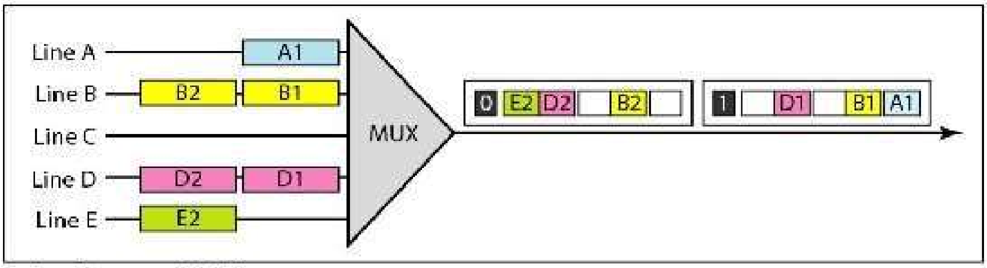
In this method, if we have "n" input lines, the frame contains a fixed number of time slots (at least "n" time slot). In an asynchronous system, if we have "n" input lines, the frame contains no more than "n" slots, with m<n where "m" is frame and "n" is no of slots.

In this way, asynchronous TDM supports the same number of input lines as synchronous with lower capacity link. It can support more devices than synchronous TDM

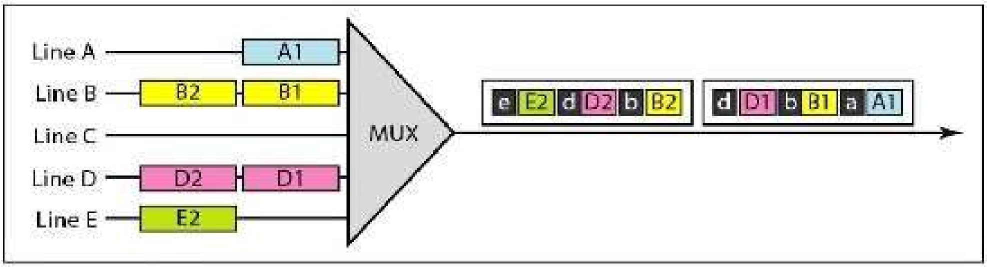




b. Statistical TDM



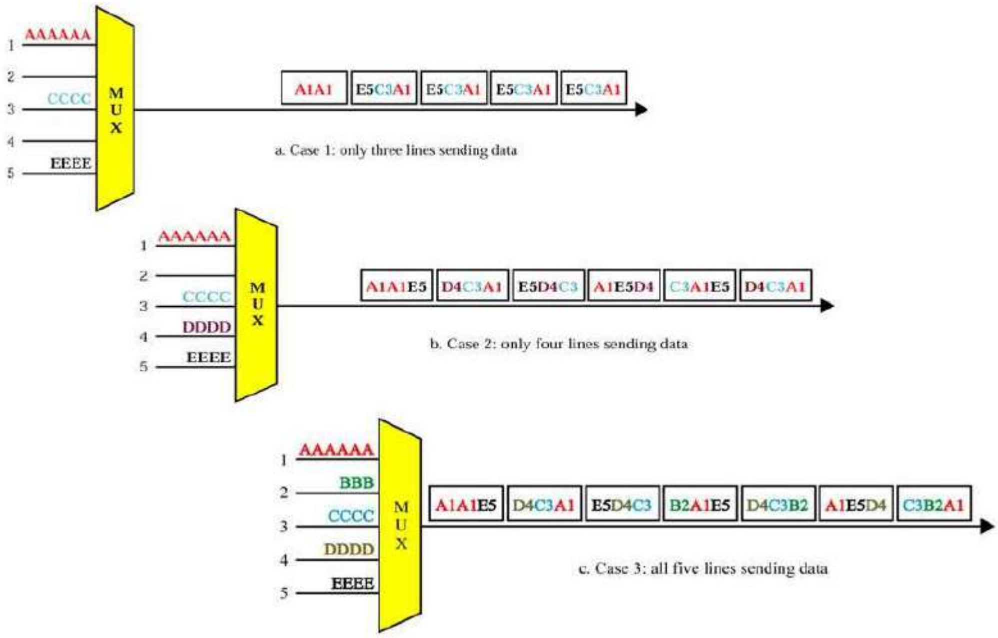
a. Synchronous TDM



b. Statistical TDM

The number of time slots in an asynchronous TDM frame (m) is based on a statistical analysiss of the number of inputs of input lines, that are likely to be transmitting at any given time. The multiplexer scans the input lines, accepts portions of data until a frame is filled, and then sends the frame across the line.

If there are not enough data to fill all the slots in a frame, the frame is transmitted only partially filled. So full-link capacity may not be used 100% of the time.



a. Case 1: only three lines sending data

b. Case 2: only four lines sending data

c. Case 3: all five lines sending data

In the above example, the frame size is three slots. In the 1st case, only 3 of the 5 devices have data to send. In the 2nd case, only 4 devices are sending data, one more than the number of slots per frame. In the 3rd case, all the lines are sending data. In each case, the multiplexer scans the devices in order from 1 to 5, filling time slots as it encounters data to be sent.

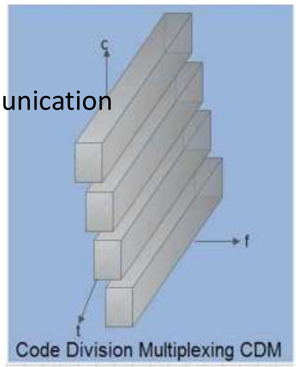
Code Division Multiplexing

- 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 travels with these codes independently travelling inside the whole bandwidth. The receiver in this case, knows in advance chip code signal it has to receive signals.
- CDM is widely used in so-called second-generation (2G) and third-generation 3G wireless communications. The technology is used in ultra-high-frequency (UHF) cellular telephone systems in the 800-MHz and 1.9-GHz bands. This is a combination of analog- to-digital conversion and spread spectrum technology.
- CDM may be defined as a form of multiplexing where the transmitter encodes the signal using a pseudo-random sequence. CDM involves the original digital signal with a spreading code. This spreading has the effect of spreading the spectrum of the signal greatly and reducing the power over anyone part of the spectrum. On the other hand, the

  decode the received signal. Each

  different random sequence corresponds to a different communication

  channel from multiple stations.



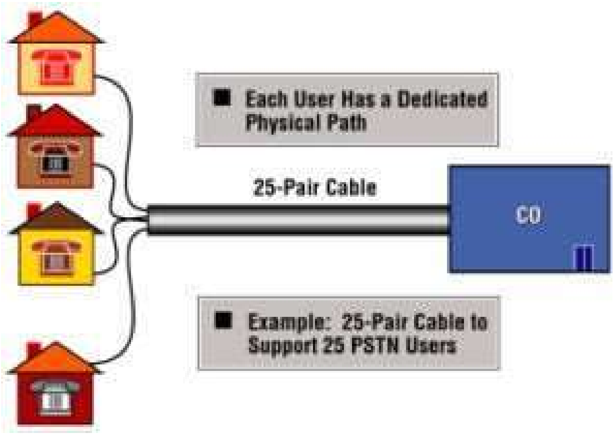
Code Division Multiplexing CDM

- ➢ Code Division Multiplexing assigns each channel its own code to make them separate from each other. These unique underlying codes, which ~hen decoded restore' the original removing the effect of the other coded channels.

  Guard spaces are realized by using codes with orthogonal codes..Figure explains how all channels C_i, use the same frequency at the same time for transmission.
- It may be understood that a single bit may be transmitted by modulating a series of signal elements at different frequencies in some particular order. These numbers of different frequencies per bit are called as the chip rate. If one or more bits are transmitted at the same frequency, it is called as frequency hopping. This will happen only when the chip rate ,is less than one because chip rate is the ratio of frequency and bit. At the receiving side, receiver decodes a 0 or a 1 bit by checking these frequencies in the correct order.

Space-Division Multiplexing

- When we want to transmit multiple messages, the goal is maximum reuse of the given resources: time and frequency. Time-Division Multiplexing (TDM), operates by dividing the time up into time slices, so that the available time can be reused. Frequency-Division Multiplexing (FDM), operates by dividing up the frequency into transmission bands, so that the frequency spectrum can be reused.


■ Each User Has a Dedicated Physical Path
25-Pair Cable
CO
■ Example: 25-Pair Cable to Support 25 PSTN Users

  However, if we remember our work with directional antennas, we can actually reuse both time and frequency, by transmitting our information along parallel channels. This is known as **Space-Division Multiplexing**.

- Space division multiplexing (SDM) is nothing more than the provision of multiple fixed bandwidth channels by multiple physical paths (i.e., pairs of wires or optical fibers). A good example of SDM is the use of a 25-pair cable to carry the conversations of 25 individual users from the customer's premises to the local telephone company’s central office location.
- SDM is not the most efficient technique from the standpoint of outside plant resources, but it does play a role in all carrier networks. A given copper or fiber facility has a finite capacity for information. When that capacity is exhausted, SDM is the onlyalternative.
- By some arguments, SDM is not a multiplexing scheme at all, since it does not support multiple communication channels on a single medium. However, the concept is an important one because it occurs both in the deployment of transmission facilities as well as the internal architecture of some switches.

UNT-5 Switching & Routing

5.1 Circuit Switching networks

5.2 Packet Switching principles

5.3 X.25

5.4 Routing in Packet switching

5.5 Congestion

5.6 Effects of congestion, congestion control

5.7 Traffic Management

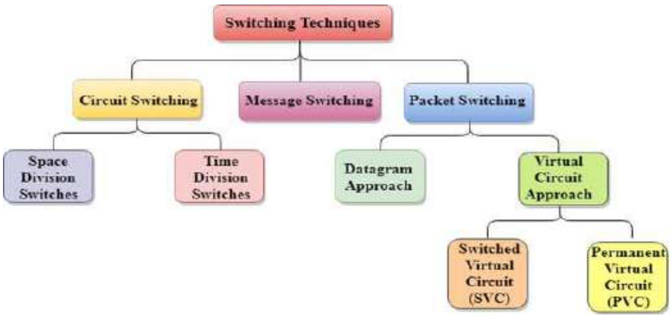
5.8 Congestion Control in Packet Switching Network

5.1 Switching techniques

In large networks, there can be multiple paths from sender to receiver. The switching technique will decide the best route for data transmission.

Switching technique is used to connect the systems for making one-to-one communication.

Classification Of Switching Techniques

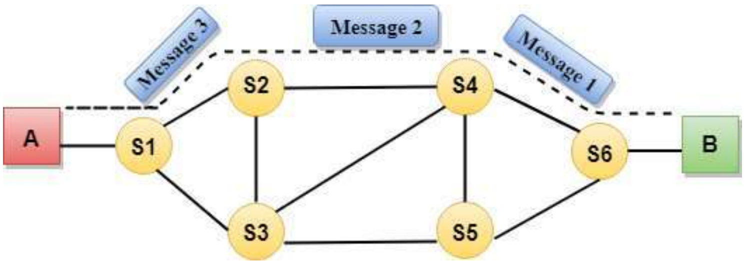


Circuit Switching

- Circuit switching is a switching technique that establishes a dedicated path between sender and receiver.
- In the Circuit Switching Technique, once the connection is established then the dedicated path will remain to exist until the connection is terminated.
- Circuit switching in a network operates in a similar way as the telephone works.
- A complete end-to-end path must exist before the communication takes place.
- In case of circuit switching technique, when any user wants to send the data, voice, video, a request signal is sent to the receiver then the receiver sends back the acknowledgment to ensure the availability of the dedicated path. After receiving the acknowledgment, dedicated path transfers the data.
- Circuit switching is used in public telephone network. It is used for voice transmission.
- Fixed data can be transferred at a time in circuit switching technology.

Communication through circuit switching has 3 phases:

- Circuit establishment
- Data transfer
- Circuit Disconnect



Circuit Switching can use either of the two technologies:

Space Division Switches:

- Space Division Switching is a circuit switching technology in which a single transmission path is accomplished in a switch by using a physically separate set of crosspoints.
- Space Division Switching can be achieved by using crossbar switch. A crossbar switch is a metallic crosspoint or semiconductor gate that can be enabled or disabled by a control unit.
- The Crossbar switch is made by using the semiconductor. For example, Xilinx crossbar switch using FPGAs.
- Space Division Switching has high speed, high capacity, and nonblocking switches.

Space Division Switches can be categorized in two ways:

- **Crossbar Switch**
- **Multistage Switch**

Crossbar Switch

The Crossbar switch is a switch that has n input lines and n output lines. The crossbar switch has n^2 intersection points known as **crosspoints.**

Disadvantage of Crossbar switch:

The number of crosspoints increases as the number of stations is increased. Therefore, it becomes very expensive for a large switch. The solution to this is to use a multistage switch.

Multistage Switch

- Multistage Switch is made by splitting the crossbar switch into the smaller units and then interconnecting them.
- It reduces the number of crosspoints.
- If one path fails, then there will be an availability of another path.

Advantages Of Circuit Switching:

- In the case of Circuit Switching technique, the communication channel is dedicated.
- It has fixed bandwidth.

Disadvantages Of Circuit Switching:

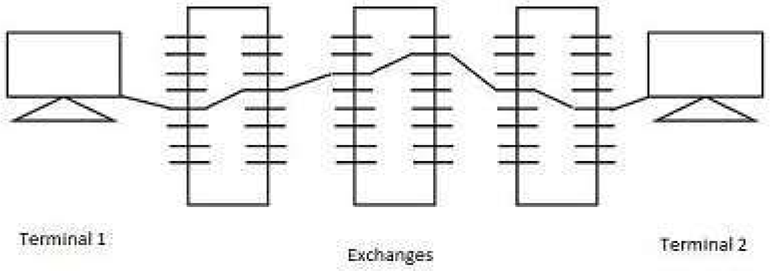
- Once the dedicated path is established, the only delay occurs in the speed of data transmission.
- It takes a long time to establish a connection approx 10 seconds during which no data can be transmitted.
- It is more expensive than other switching techniques as a dedicated path is required for each connection.
- It is inefficient to use because once the path is established and no data is transferred, then the capacity of the path is wasted.
- In this case, the connection is dedicated therefore no other data can be transferred even if the channel is free.

Time Division Switching

Time division switching comes under digital switching techniques, where the Pulse Code Modulated signals are mostly present at the input and the output ports. A digital Switching system is one, where the inputs of any PCM highway can be connected to the outputs of any PCM highway, to establish a call.

The incoming and outgoing signals when received and re-transmitted in a different time slot, is called **Time Division Switching.** The digitized speech information is sliced into a sequence of time intervals or slots. Additional voice circuit slots, corresponding to other users are inserted into this bit stream of data. Hence, the data is sent in time frames.

The main difference between space division multiplexing and time division multiplexing is sharing of Crosspoints. Crosspoints are not shared in space division switching, whereas they can be shared in time division multiplexing, for shorter periods. This helps in reassigning the Crosspoints and its associated circuitry for other connections as well.



Time division switches use time division multiplexing, in switching. The two popular methods of TDM are TSI (Time and Slot Interchange) and TDM bus. The data sent at the transmitter reaches the receiver in the same order, in an ordinary time division multiplexing whereas, in TSI mechanism, the data sent is changed according to the ordering of slots based on the desired connections. It consists of RAM with several memory locations such as input, output locations and control unit.

Both of the techniques are used in digital transmission. The TDM bus utilizes multiplexing to place all the signals on a common transmission path. The bus must have higher data rate than individual I/O lines. The main advantage of time division multiplexing is that, there is no need of Crosspoints. However, processing each connection creates delay as each time slot must be stored by RAM, then retrieved and then passed on.

Message Switching

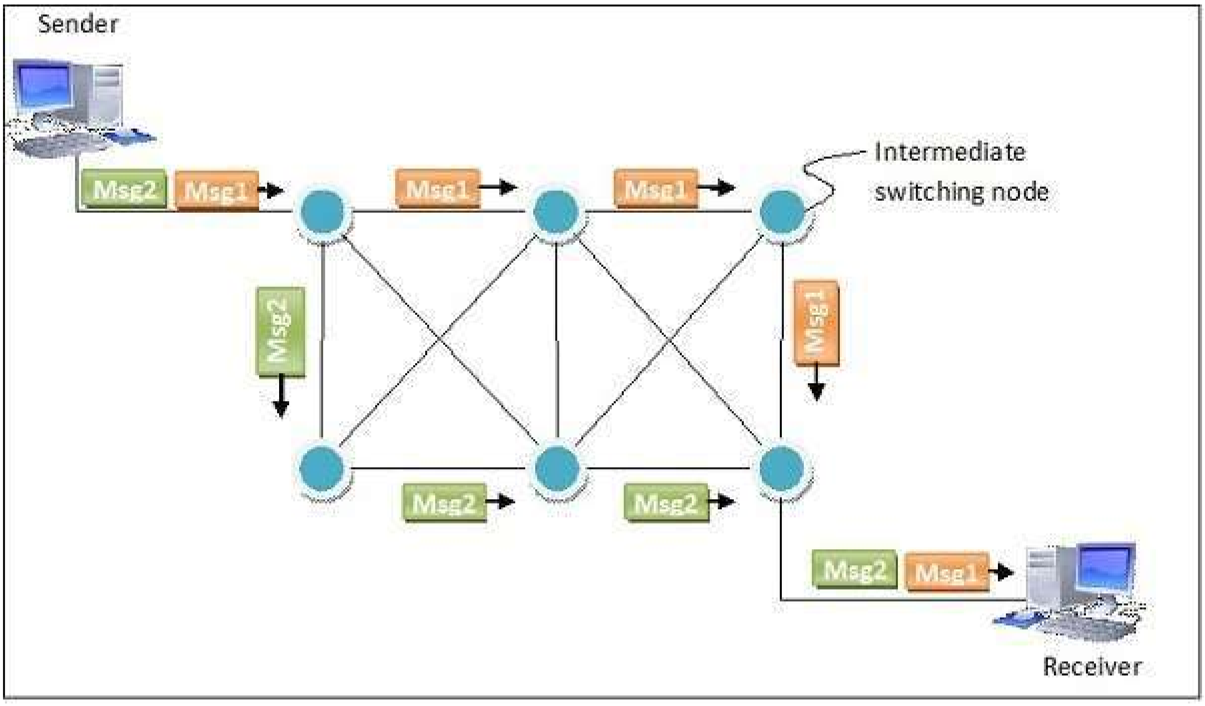
Message switching is a connectionless network switching technique where the entire message is routed from the source to the destination, at a time. It was a precursor of packet switching.

During message routing, every intermediate switch in the network stores the whole message. If the entire network's resources are engaged or the network becomes blocked, the message-switched network stores and delays the message until ample resources become available for effective transmission of the message.

Process

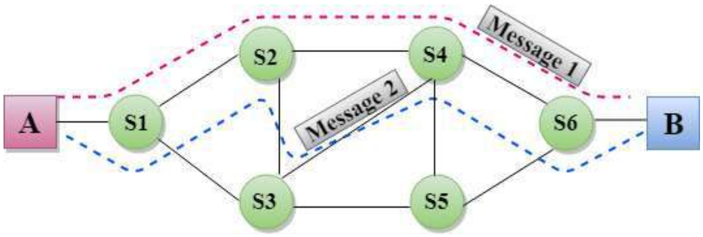
Packet switching treats each message as an individual unit. Before sending the message, the sender node adds the destination address to the message. Then it is delivered entirely to the next intermediate switching node. The intermediate node stores the message in its entirety, checks for transmission errors, inspects the destination address and then delivers it to the next node. The process continues till the message reaches the destination. The incoming message is not discarded if the required outgoing circuit is busy. It is stored in a queue for that route and retransmitted when the required route is available. This is called store and forward network.

The following diagram represents routing of two separate messages from the same source to same destination via different routes, using message switching –



- Message Switching is a switching technique in which a message is transferred as a complete unit and routed through intermediate nodes at which it is stored and forwarded.
- In Message Switching technique, there is no establishment of a dedicated path between the sender and receiver.

- The destination address is appended to the message. Message Switching provides a dynamic routing as the message is routed through the intermediate nodes based on the information available in the message.
- Message switches are programmed in such a way so that they can provide the most efficient routes.
- Each and every node stores the entire message and then forward it to the next node. This type of network is known as **store and forward network.**
- Message switching treats each message as an independent entity.



Advantages Of Message Switching

- Data channels are shared among the communicating devices that improve the efficiency of using available bandwidth.
- Traffic congestion can be reduced because the message is temporarily stored in the nodes.
- Message priority can be used to manage the network.
- The size of the message which is sent over the network can be varied. Therefore, it supports the data of unlimited size.
- Broadcasting messages requires much less bandwidth than circuit switching.
- It does not have to deal with out of order packets or lost packets as in packet switching.

Disadvantages Of Message Switching

- The message switches must be equipped with sufficient storage to enable them to store the messages until the message is forwarded.
- The Long delay can occur due to the storing and forwarding facility provided by the message switching technique.

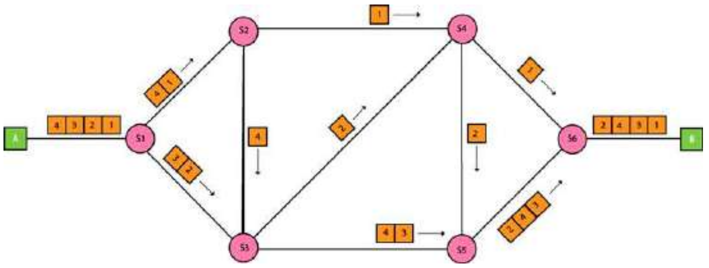
5.2 Packet Switching

Message switching is a connectionless network switching technique where the entire message is routed from the source to the destination, at a time. It was a precursor of packet switching. During message routing, every intermediate switch in the network stores the whole message. If the entire network's resources are engaged or the network becomes blocked, the message-switched network stores and delays the message until ample resources become available for effective transmission of the message.

Process

Packet switching treats each message as an individual unit. Before sending the message, the sender node adds the destination address to the message. Then it is delivered entirely to the next intermediate switching node. The intermediate node stores the message in its entirety, checks for transmission errors, inspects the destination address and then delivers it to the next node. The process continues till the message reaches the destination. The incoming message is not discarded if the required outgoing circuit is busy. It is stored in a queue for that route and retransmitted when the required route is available. This is called store and forward network.

- Packet switching is a digital network transmission process in which data is broken into suitably-sized pieces or blocks for fast and efficient transfer via different network devices.
- When a computer attempts to send a file to another computer, the file is broken into packets so that it can be sent across the network in the most efficient way.
- These packets are then routed by network devices to the destination.
- The packet switching is a switching technique in which the message is sent in one go, but it is divided into smaller pieces, and they are sent individually.
- The message splits into smaller pieces known as packets and packets are given a unique number to identify their order at the receiving end.
- Every packet contains some information in its headers such as source address, destination address and sequence number.
- Packets will travel across the network, taking the shortest path as possible.
- All the packets are reassembled at the receiving end in correct order.
- If any packet is missing or corrupted, then the message will be sent to resend the message.
- If the correct order of the packets is reached, then the acknowledgment message will be sent.



Approaches Of Packet Switching:

There are two approaches to Packet Switching:

A) Connectionless Packet Switching (Datagram Packet Switching)

B) Connection-Oriented Packet Switching (Virtual Circuit Switching)

A) Connectionless Packet switching (Datagram) :

Each packet contains complete addressing or routing information and is routed individually. This can result in out-of-order delivery and different paths of transmission, depending on the variable loads on different network nodes (adapters, switches and routers) at any given time. Also known as datagram switching.

In connectionless packet switching, each packet has the following information written in its header section:

- The destination address
- The source address
- Total number of pieces
- The sequence number (Seq#) needed to enable reassembly

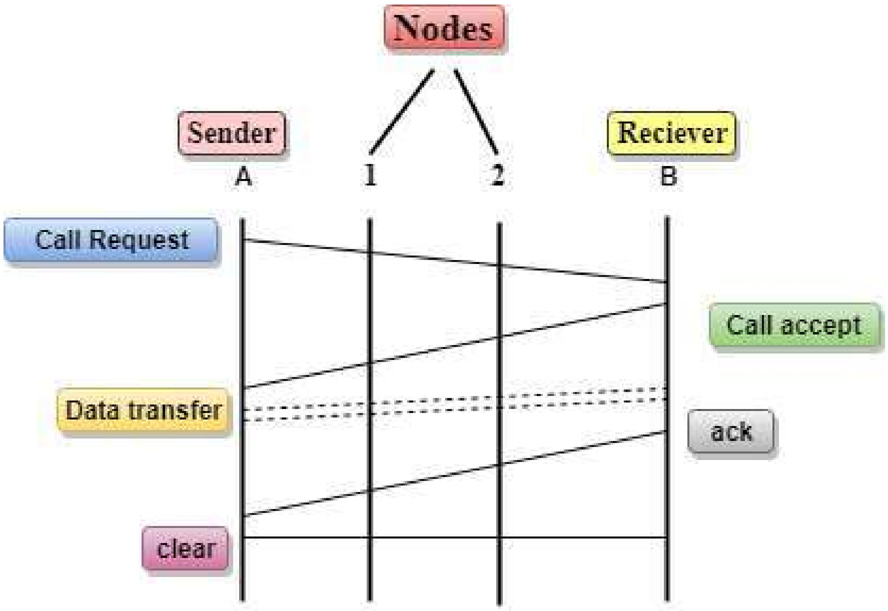
After reaching the destination through different routes, the packets are rearranged to form the original message.

- It is a packet switching technology in which packet is known as a datagram, is considered as an independent entity. Each packet contains the information about the destination and switch uses this information to forward the packet to the correct destination.
- The packets are reassembled at the receiving end in correct order.
- In Datagram Packet Switching technique, the path is not fixed.
- Intermediate nodes take the routing decisions to forward the packets.
- Datagram Packet Switching is also known as connectionless switching.

B) Connection Oriented Circuit Switching (Virtual):

- Virtual Circuit Switching is also known as connection-oriented switching.
- In the case of Virtual circuit switching, a preplanned route is established before the messages are sent.
- Call request and call accept packets are used to establish the connection between sender and receiver.
- In this case, the path is fixed for the duration of a logical connection.
- The circuit made between two stations at the start of communication using packet switching service is referred as Virtual Circuit. Virtual circuit can have bi-directional communication path between two DTEs (Data Terminal Equiments) or between DTE and DCE (Data circuit-terminating equipment). Each of the paths in virtual circuits are identified by unique DLCI (Data Link Connection Identifier) field.
- More than one virtual circuits are multiplexed on single physical circuit connection for transmission over network. This reduces network complexity and cost of network equipments.

Concept of virtual circuit switching through a diagram:



- In the above diagram, A and B are the sender and receiver respectively. 1 and 2 are the nodes.
- Call request and call accept packets are used to establish a connection between the sender and receiver.
- When a route is established, data will be transferred.
- After transmission of data, an acknowledgment signal is sent by the receiver that the message has been received.
- If the user wants to terminate the connection, a clear signal is sent for the termination.

Virtual Circuit Switching Types:

a) **Switched Virtual Circuit(SVC)**

b) **Permanent Virtual Circuit (PVC)**

A) SVC-Switched Virtual Circuit

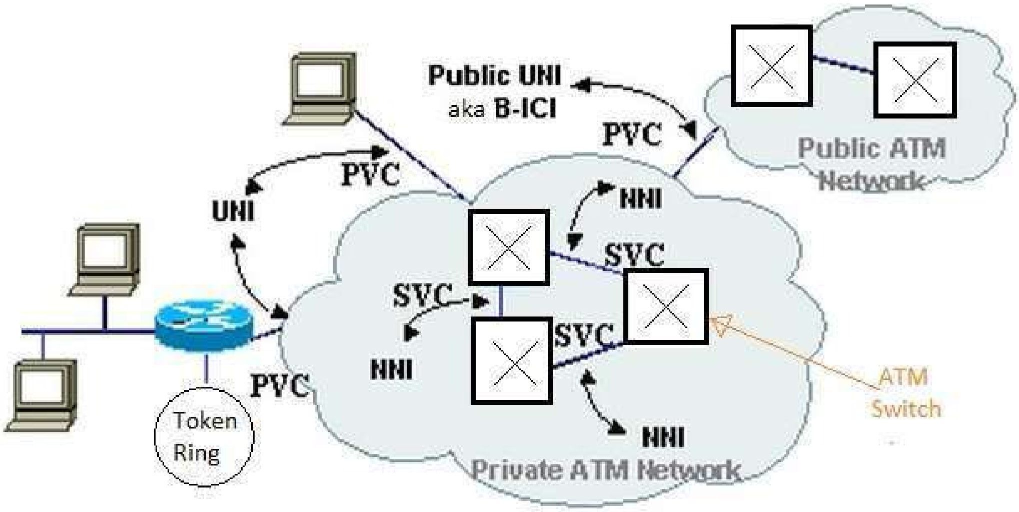
• Switched Virtual Circuits (SVCs) are temporary connections created for the purpose of information transfer. There are four steps to establish SVC connection viz. call setup, data transfer, Idle and call termination.

• Once the virtual circuit is cleared, it has to be re-establish in order to transmit any further data.

- As the circuit is not fixed (i.e. not open) all the time, the cost to use the SVC service is less. It is established on need basis in order to transmit data as required.

B) PVC-Permanent Virtual Circuit

- Permanent Virtual Circuits (PVCs) are permanent connections established for the sole purpose of frequent as well as consistent data transfer.
- As it is like leased line, PVC connection do not require to be be established using call setup or termination states. PVC will be either in data transfer mode or in IDle mode.
- In data transfer mode, data is transmitted between two DTE devices over virtual circuit path.
- In Idle mode, connection between DTEs is available but data transfer is not progressing.
- Unlike Switched Virtual Circuit, PVCs are not terminated during idle state.
- As this type of virtual circuit connection is permanent, data transfer can take place as soon as it is ready to transmit.



SVC (Switched Virtual Circuit) vs PVC (Permanent Virtual Circuit)

Differences b/w DATAGRAM APPROACH and VIRTUAL CIRCUIT approach

Datagram approach	Virtual Circuit approach
Node takes routing decisions to forward the packets.	Node does not take any routing decision.
Congestion cannot occur as all the packets travel in different directions.	Congestion can occur when the node is busy, and it does not allow other packets to pass through.

It is more flexible as all the packets are treated as an independent entity.	It is not very flexible.

Advantages Of Packet Switching:

- **Cost-effective:** In packet switching technique, switching devices do not require massive secondary storage to store the packets, so cost is minimized to some extent. Therefore, we can say that the packet switching technique is a cost-effective technique.
- **Reliable:** If any node is busy, then the packets can be rerouted. This ensures that the Packet Switching technique provides reliable communication.
- **Efficient:** Packet Switching is an efficient technique. It does not require any established path prior to the transmission, and many users can use the same communication channel simultaneously, hence makes use of available bandwidth very efficiently.

Disadvantages Of Packet Switching:

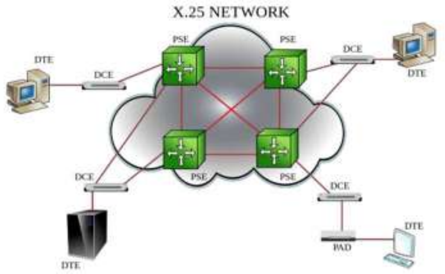
- Packet Switching technique cannot be implemented in those applications that require low delay and high-quality services.
- The protocols used in a packet switching technique are very complex and requires high implementation cost.
- If the network is overloaded or corrupted, then it requires retransmission of lost packets. It can also lead to the loss of critical information if errors are nor recovered.

X.25

X. 25 is the name given to **a suite of protocols used for packet-switched wide area network communication**. Defined by the International Telegraph and Telephone Consultative Committee in 1976, X. 25 had the original purpose of carrying voice signals over analog telephone lines.

- X.25 is an ITU-T standard protocol suite for packet switched wide area network (WAN) communication.
- X.25 was originally defined by the International Telegraph and Telephone Consultative Committee (CCITT, now ITU-T) in a series of drafts and finalized in a publication known as The Orange Book in 1976.
- .X.25 is a family of protocols that was popular during the 1980s with telecommunications companies and in financial transaction systems such as automated teller machines.
- X.25 is a standard suite of protocols used for packet switching across computer networks. The X.25 protocols works at the physical, data link, and network layers (Layers 1 to 3) of the OSI model.
- *Each X.25 packets contains up to 128 bytes of data. The X.25 network handles packet assembly at the source device, delivery, and then dis-assembly at the destination. X.25*

packet delivery technology includes not only switching and network-layer routing, but also error checking and re-transmission logic should delivery failures occur. X.25 supports multiple simultaneous conversations by multiplexing packets and using virtual communication channels.



- Based upon existing analog copper lines that experience a high number of errors
- Uses the virtual circuit approach
- An X.25 WAN consists of packet-switching exchange (PSE) nodes as the networking hardware, and leased lines, plain old telephone service connections or ISDN connections as physical links
- Provides a way to send packets across a packet-switched public data network
- The redundant error checking is done at each node
- X.25 was originally designed more than 25 years ago to carry voice over analog telephone lines (dialup networks). Typical applications of X.25 today include automatic teller machine networks and credit card verification networks. X.25 also supports a variety of mainframe terminal/server applications.
- With the widespread acceptance of Internet Protocol (IP) as a standard for corporate networks, many X.25 applications are now being migrated to cheaper solutions using IP as the network layer protocol and replacing the lower layers of X.25 with Ethernet or ATM hardware.

Architecture

- The X.25 specification defines only the interface between a subscriber (DTE) and an X.25 network (DCE). X.75, a very similar protocol to X.25, defines the interface between two X.25 networks to allow connections to traverse two or more networks.
- X.25 originally defined three basic protocol levels or architectural layers. The layer numbers were dropped to avoid confusion with the OSI Model layers.

Physical layer

- This layer specifies the physical, electrical, functional and procedural characteristics to control the physical link between a DTE and a DCE.

- Common implementations use X.21, EIA-232, EIA-449 or other serial protocols.
- It lays out the physical, electrical and functional characteristics that interface between the computer terminal and the link to the packet switched node. X.21 physical implementer is commonly used for the linking

Data link layer

- The data link layer consists of the link access procedure for data interchange on the link between a DTE and a DCE.
- In its implementation, the, link accessed procedure balanced (lapb) is a data link protocol that manages a communication session and controls the packet framing.
- It is a bit-oriented protocol that provides error correction and orderly delivery.
- It comprises the link access procedures for exchanging data over the link. Here, control information for transmission over the link is attached to the packets from the packet layer to form the LAPB frame (Link Access Procedure Balanced). This service ensures a bit-oriented, error-free, and ordered delivery of frames.

Packet layer

- This layer defined a packet-layer protocol for exchanging control and user data packets to form a packet-switching network based on virtual calls, according to the packet layer.
- X.25 provides a set of user facilities defined and described in ITU-T Recommendation

  X.2. The X.2 user facilities fall into five categories:

  1. Essential facilities;
  2. Additional facilities;
  3. Conditional facilities;
  4. Mandatory facilities.
  5. Optional facilities.

- This layer defines the format of data packets and the procedures for control and transmission of the data packets. It provides external virtual circuit service. Virtual circuits may be of two types: virtual call and permanent virtual circuit. The virtual call is established dynamically when needed through call set up procedure, and the circuit is relinquished through call clearing procedure. Permanent virtual circuit, on the other hand, is fixed and network assigned.

Advantages of X.25

- Frame delivery is more reliable
- Frames are delivered in order
- Retransmission of frames is possible
- Flow control is provided
- X.25 supports the switched virtual circuits and permanent circuits

Disdvantage of X.25

- X.25 is much slower than Frame relay

5.5 Network Congestion

Just like in road congestion, Network Congestion occurs when a network is not able to adequately handle the traffic flowing through it. Network congestion is usually a temporary state of a network rather than a permanent feature.

Network congestion in data networking is the reduced quality of service that occurs when a network node or link is carrying more data than it can handle. Its effects include queueing delay, packet loss or the blocking of new connections.

In this section, we will discuss five (5) common causes of network congestion including:

- Over-subscription
- Poor network design/mis-configuration
- Over-utilized devices
- Faulty devices
- Security attack

Over-Subscription

Over-Subscription where a system (e.g. a network) is handling more traffic than it was designed to handle per time. Over-subscription is usually done on purpose as it may result in cost savings.

For example, An organization has 100 users and it has been determined that a 100Mbps Internet link will be suitable for all these users.

Now imagine that most of the staff of this organization work from home. In this case, it will be more cost efficient to go for a lower link capacity, say 50Mbps, since only a handful of employees will be using the link per time. But what happens when there is a company-wide meeting and all employees come into the office? You guessed right – Network congestion.

Poor Network Design/Mis-Configuration

A more serious cause of network congestion is poor design or device Mis-Configuration. Take for example a broadcast storm, where a large volume of broadcast and/or multicast traffic is seen on the network within a short time, resulting in severe performance degradation.

Since broadcasts are contained within subnets, the larger the subnet the more serious the effect of a broadcast storm. Therefore, a network that has been designed with large subnets without giving proper consideration to broadcast storms can result in network congestion.

Over-Utilized Devices

Devices such as routers, switches, and firewalls have been designed to handle certain network throughput.

For example, the Juniper MX5 has a capacity of 20Gbps. Therefore, constantly pushing ~20Gbps of traffic through that device means that the device will be over-utilized and will likely result in high CPU utilization and packet drops, leading to congestion on the network.

Another issue related to over-utilized devices that can cause network congestion is Bottlenecks. As in most hierarchical designs where multiple devices feed into a higher-level device, care must be taken to ensure that the higher-level device is capable of handling all the traffic from the lower-level devices.

Faulty Devices

Example (lower speed device): Network performance assessment for an organization. They were buying 100Mbps link capacity from their ISP but the users on the network were struggling to connect to the Internet effectively.

They complained that the network was always “slow” (user speak for network congestion) even when few people were on the network. Upon investigation, it was found that, their ISP was truly giving the agreed upon 100Mbps, the edge device was only providing 30Mbps to the network!

Apart from the fact that this organization had wrongly terminated the link on a FastEthernet interface (which gives a theoretical speed of 100Mbps but much lower practical speed), that interface was also faulty. By moving the ISP link to another interface (we used a GigabitEthernet interface instead), the performance problem was solved.

Security Attack

Example (attacker using server) : In another organization, a network of about 10 users had poor browsing experience even with the 4Mbps link they were getting from their ISP.

Ideally, this capacity should have been enough because the users were not doing anything heavy on the Internet – just emails, web searches, and normal user activities.

Upon investigation, it was discovered that one of their servers had been compromised and it seems the attacker was using this server to host illicit content resulting in a huge amount of traffic being sent to/from this server. By cleaning up this server, the congested network was once again “free” for normal user traffic.

Other security attacks that can result in network congestion include viruses, worms, and Denial of Service (DoS) attacks.

5.6 Effects of Network Congestion

Everyone on a network generally “feels” the effects of network congestion. They may not be able to explain it in technical terms but will say things like “The connection is so slow”, “I can’t open web pages”, “The network is really bad, I can’t hear you”.

From a technical perspective, the effects of a congested network include:

- **Delay:**

  Also known as Latency, Delay is the time it takes for a destination to receive the packet sent by the sender. For example, the time it takes for a webpage to load is a result of how long it takes for the packets from the web server to get to the client. Another evidence of delay is the buffering you experience when watching a video, say on YouTube.

- **Packet Loss:**

  While packets may take a while to get to their destination (delay), packet loss is an even more negative effect of network congestion. This is especially troubling for applications like Voice over IP (VoIP) that do not deal well with delay and packet loss, resulting in dropped calls and Call Detail Records, lag, robotic voices, and so on.

- **Timeouts:**

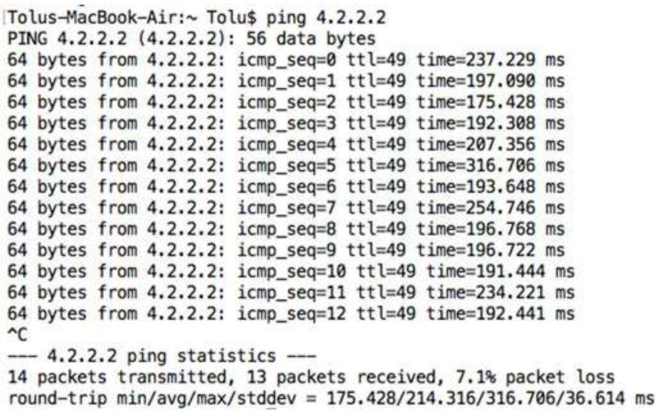
  Network congestion can also result in timeouts in various applications. Since most connections will not stay up indefinitely waiting for packets to arrive, this can result in lost connections.

5.6 Troubleshooting Network Congestion

Feeling the effects of network congestion is one thing but actually confirming that a network is congested is another. In this section, we will look at some activities that can be performed to confirm the congestion of a network.

1. Ping

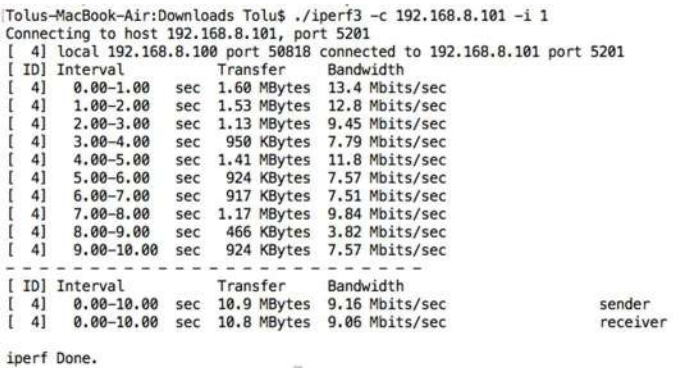
One of the fastest ways to check if a network is congested is to use Ping because not only can it detect packet loss, it can also reveal delay in a network i.e. through the round-trip time (RTT). Using a tool like MTR (which combines ping and traceroute) can also reveal parts of the network where congestion is occurring.



```
Tolus-MacBook-Air:~ Tolu$ ping 4.2.2.2
PING 4.2.2.2 (4.2.2.2): 56 data bytes
64 bytes from 4.2.2.2: icmp_seq=0 ttl=49 time=237.229 ms
64 bytes from 4.2.2.2: icmp_seq=1 ttl=49 time=197.090 ms
64 bytes from 4.2.2.2: icmp_seq=2 ttl=49 time=175.428 ms
64 bytes from 4.2.2.2: icmp_seq=3 ttl=49 time=192.308 ms
64 bytes from 4.2.2.2: icmp_seq=4 ttl=49 time=207.356 ms
64 bytes from 4.2.2.2: icmp_seq=5 ttl=49 time=316.706 ms
64 bytes from 4.2.2.2: icmp_seq=6 ttl=49 time=193.648 ms
64 bytes from 4.2.2.2: icmp_seq=7 ttl=49 time=254.746 ms
64 bytes from 4.2.2.2: icmp_seq=8 ttl=49 time=196.768 ms
64 bytes from 4.2.2.2: icmp_seq=9 ttl=49 time=196.722 ms
64 bytes from 4.2.2.2: icmp_seq=10 ttl=49 time=191.444 ms
64 bytes from 4.2.2.2: icmp_seq=11 ttl=49 time=234.221 ms
64 bytes from 4.2.2.2: icmp_seq=12 ttl=49 time=192.441 ms
^C
--- 4.2.2.2 ping statistics ---
14 packets transmitted, 13 packets received, 7.1% packet loss
round-trip min/avg/max/stddev = 175.428/214.316/316.706/36.614 ms
```

2. LAN Performance Tests

A tool like iPerf can be very useful in determining performance issues on a network, measuring statistics like bandwidth, delay, jitter, and packet loss. This can help reveal bottlenecks on the network and also identify any faulty devices/interfaces.

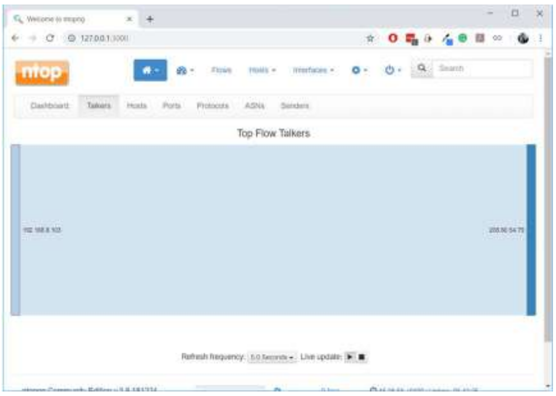


```
Tolus-MacBook-Air:Downloads Tolu$ ./iperf3 -c 192.168.8.101 -i 1
Connecting to host 192.168.8.101, port 5201
[  4] local 192.168.8.100 port 50818 connected to 192.168.8.101 port 5201
[ ID] Interval           Transfer     Bandwidth
[  4]   0.00-1.00   sec  1.60 MBytes  13.4 Mbits/sec
[  4]   1.00-2.00   sec  1.53 MBytes  12.8 Mbits/sec
[  4]   2.00-3.00   sec  1.13 MBytes  9.45 Mbits/sec
[  4]   3.00-4.00   sec   950 KBytes  7.79 Mbits/sec
[  4]   4.00-5.00   sec  1.41 MBytes  11.8 Mbits/sec
[  4]   5.00-6.00   sec   924 KBytes  7.57 Mbits/sec
[  4]   6.00-7.00   sec   917 KBytes  7.51 Mbits/sec
[  4]   7.00-8.00   sec  1.17 MBytes  9.84 Mbits/sec
[  4]   8.00-9.00   sec   466 KBytes  3.82 Mbits/sec
[  4]   9.00-10.00  sec   924 KBytes  7.57 Mbits/sec
- - - - - - - - - - - - - - - - - - - - - - - - -
[ ID] Interval           Transfer     Bandwidth
[  4]   0.00-10.00  sec  10.9 MBytes  9.16 Mbits/sec                  sender
[  4]   0.00-10.00  sec  10.8 MBytes  9.06 Mbits/sec                  receiver

iperf Done.
```

3. Bandwidth Monitoring

During the investigation of the compromised server I mentioned above, we used a tool called ntopng to discover “Top Talkers” which revealed that the server was using up all the bandwidth on the network. In the same way, tools that monitor bandwidth can reveal network congestion especially during a security attack or if a particular host is using up all the bandwidth.



5.8 Congestion Control in Packet Switching Network

Different congestion control approaches in packet switching network (datagram subnet and also in virtual circuit subnets) are :

i) Choke packets
ii) Load Shedding
iii) Jitter control

i) Choke Packets :

In this technique, each router associates a real variable with each of its output lines. This real value “u” has a value between 0 and 1 and it indicates the percentage utilization of that line. If he value “u” goes above the threshold (defined values that determine if a statistic is above, below, or within a normal range on your network) then the output line will enter into a warning state. The router will check each newly arriving packet to see if its output line is in the warning state. If it is in the warning state, then the router will send back a choke packet signal to the sending host. Then the sender will not generate any more choke packets. Depending on the threshold vale, the choke packets can contain a mild warning, a stern warning or an ultimatum.

Drawback :

The action to be taken by the source host on receiving a choke packet is voluntary and not compulsory.

ii) Load Shedding

It is one of the simplest and more effective techniques. In this method, whenever a router finds that there is congestion in the network, it simply starts dropping out the packets.

The principle of load shedding states that, when the routers are being inundated (overwhelmed) by the packets that they cant not handle, they should simply throw the packets away. A router which is flooding with packets due to congestion can drop any packet randomly.

The policy for dropping a packet depends on the type of packet. For file transfer, the old packet is more important than newer one and for multimedia, new packet is more important than older one.

An intelligent discard policy can be decide depending on the application. There are various effective ways which requires cooperation from the sender. For many applications, some packets are more important than others. So, sender can mark the packets in priority classes to indicate how important they are. If such a priority policy is implemented than intermediate nodes can drop packets from the lower priority classes and use the available bandwidth for the more important packets

iii) Jitter Control

Jitter is defined as the variation in delay for the packets belonging to the same flow. The real time audio and video cannot tolerate jitter on the other hand the jitter does not matter if the packets are carrying an information contained in a file. For the audio and video transmission if the packets take 20 msec to 30msec to reach the destination, it does not matter, provided that the delay remains constant. The quality of sound or video wll be hampered if the delays associated with different packets have different values

When a packet arrives at a router, the router will check to see whether the packet is behind or ahead and by what time. This information is stored in the packet and updated every hop. If the packet is ahead of the schedule then the router will hold it for slightly longer time and if the packet is behind the schedule, then the router will try to send it out as quickly as possible.

Decongesting a network

The fix for a Congested Network will Depend on the Cause:

- For oversubscribed links, you may need to purchase more bandwidth from your service provider. Some service providers also allow you to temporarily boost your bandwidth for a small fee. You may also want to implement Quality of Service (QoS) features which will ensure that even in the event of congestion, critical applications can still function.
- Layer 2 loops can be prevented by using loop prevention protocols such as Spanning Tree Protocol (STP). A poor network design can be more difficult to fix since the network is probably in use. For such cases, incremental changes can be made to improve the network and remove congestion.
- Over-Utilized devices may need to be swapped out. Alternatively, the capacity of the system can be increased by implementing high-availability features such as clustering and stacking.
- Faulty devices definitely need to be replaced. In some cases (like the example I gave above about the 100Mbps link reduced to 30Mbps), only a part of the device (e.g. an interface) needs to be replaced.
- Security attacks need to be combated as soon as they are discovered. In the case of the compromised server, the first thing we did was to remove that server from the network completely. Since this is not always a feasible solution (e.g. the compromised device is a critical server), other temporary measures such as applying access control lists to deny the offending traffic may need to be implemented.

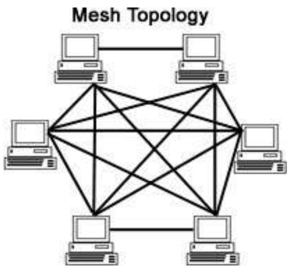
Unit- 6 LAN Technology

6.1. Topology and Transmission Media
6.2 LAN protocol architecture
6.3. Medium Access control
6.4 Bridges, Hub, Switch
6.5 Ethernet (CSMA/CD), Fiber Channel
6.6 Wireless LAN Technology

6.1 Topologies

- The way in which the connections are made of the physical devices is called the topology of the network.
- Network topology specifically refers to the physical layout of the network, especially the locations of the computers and how the cable is run between them.
- It is important to select the right topology for how the network will be used.
- Each topology has its own strengths and weaknesses.
- The four most common topologies are
  - Mesh topology
  - Bus topology
  - Star topology
  - Ring topology
  - Tree Topology
  - Hybrid Topology

Mesh topology



- A network setup where each computer and network device is interconnected with one another, allowing for most transmissions to be distributed, even if one of the connections go down.
- This topology is not commonly used for most computer networks as it is difficult and expensive to have redundant connection to every computer. However, this topology is commonly used for wireless networks.

Advantages of Mesh topology

- Data can be transmitted from different devices simultaneously. This topology can withstand high traffic.
- Even if one of the components fails there is always an alternative present. So data transfer doesn’t get affected.

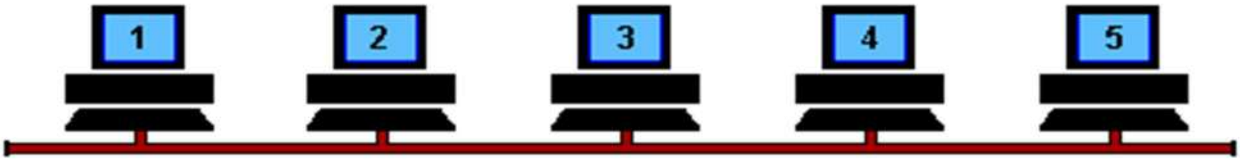
- Expansion and modification in topology can be done without disrupting other nodes.

Disadvantages of Mesh topology

- There are high chances of redundancy in many of the network connections.
- Overall cost of this network is way too high as compared to other network topologies.
- Set-up and maintenance of this topology is very difficult. Even administration of the network is tough.

Bus Topology

- The bus topology is often used when a network installation is small, simple, or temporary.
- On a typical bus network, the cable is just one or more wires, with no active electronics to amplify the signal or pass it along from computer to computer.
- This makes the bus a passive topology**.**
- When one computer sends a signal up (and down) the wire, all the computers on the network receive the information, but only one (the one with the address that matches the one encoded in the message) accepts the information. The rest disregard the message.
- Only one computer at a time can send a message; therefore, the number of computers attached to a bus network can significantly affect the speed of the network. A computer must wait until the bus is free before it can transmit.
- These factors also affect star and ring networks**.**



- Another important issue in bus networks is termination. Since the bus is a passive topology, the electrical signal from a transmitting computer is free to travel the entire length of the cable. Without termination, when the signal reaches the end of the wire, it bounces back and travels back up the wire.
- When a signal echoes back and forth along an unterminated bus, it is called ringing. To stop the signals from ringing, you attach terminators at either end of the segment. The terminators absorb the electrical energy and stop the reflections.
- Cables cannot be left unterminated in a bus network.
- *Ethernet 10Base2 (also known as thinnet) is an inexpensive network based on the bus topology.*

Advantages of Bus Topology

- There are several advantages to a bus topology:
  - The bus is simple, reliable in very small networks, easy to use, and easy to understand.
  - The bus requires the least amount of cable to connect the computers together and is therefore less expensive than other cabling arrangements.
  - It is easy to extend a bus. Two cables can be joined into one longer cable with a BNC barrel connector, making a longer cable and allowing more computers to join the network.
  - A repeater can also be used to extend a bus; a repeater boosts the signal and allows it to travel a longer distance.

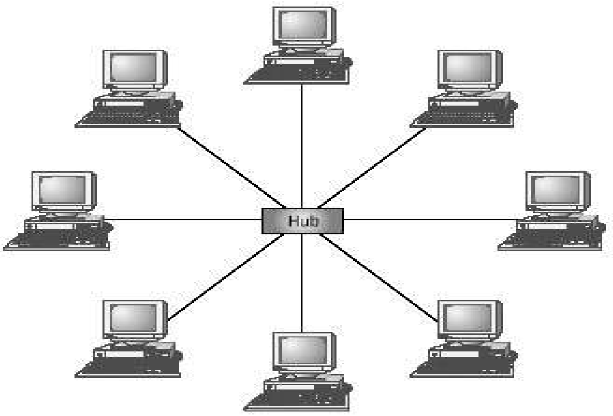
Disadvantages of Bus Topology

- Heavy network traffic can slow a bus considerably. Because any computer can transmit at any time, and computers on most bus networks do not coordinate with each other to reserve times to transmit, a bus network with a lot of computers can spend a lot of its bandwidth (capacity for transmitting information) with the computers interrupting each other instead of communicating. The problem only gets worse as more computers are added to the network.
- Each barrel connector weakens the electrical signal, and too many may prevent the signal from being correctly received all along the bus.
- It is difficult to troubleshoot a bus. A cable break or malfunctioning computer anywhere between two computers can cause them not to be able to communicate with each other. A cable break or loose connector will also cause reflections and bring down the whole network, causing all network activity to stop.

Star Topology

- In a star topology, all the cables run from the computers to a central location, where they are all connected by a device called a hub.
- Each computer on a star network communicates with a central hub that resends the message either to all the computers (in a broadcast star network) or only to the destination computer (in a switched star network). The hub in a broadcast star network can be active or passive.
- An active hub regenerates the electrical signal and sends it to all the computers connected to it. This type of hub is often called a multiport repeater. Active hubs and switches require electrical power to run.

- A passive hub, such as wiring panels or punch-down blocks, merely acts as a connection point and does not amplify or regenerate the signal.
- Passive hubs do not require electrical power to run.
- You can use several types of cable to implement a star network. A hybrid hub can accommodate several types of cable in the same star network.
- In a star topology the computers are all connected by cables to a central point.



Ring Topology

- In a ring topology, each computer is connected to the next computer, with the last one connected to the first.
- Rings are used in high-performance networks, networks that bandwidth be reserved for time-sensitive features such as video and audio, or when even performance is needed when a large number of clients access the network.
- In a ring topology computers are connected in a circle.
- Every computer is connected to the next computer in the ring, and each retransmits what it receives from the previous computer. The messages flow around the ring in one direction. Since each computer retransmits what it receives, a ring is an *active* network and is not subject to the signal loss problems a bus experiences. There is no termination because there is no end to the ring.
- Some ring networks do token passing. A short message called a token is passed around the ring until a computer wishes to send information to another computer.

- That computer modifies the token, adds an electronic address and data, and sends it around the ring.

- Each computer in sequence receives the token and the information and passes them to the next computer until either the electronic address matches the address of a computer or the token returns to its origin.
- The receiving computer returns a message to the originator indicating that the message has been received. The sending computer then creates another token and places it on the network, allowing another station to capture the token and begin transmitting. The token circulates until a station is ready to send and captures the token.
- This all happens very quickly: a token can circle a ring 200 meters in diameter at about 10,000 times a second. Some even faster networks circulate several tokens at once. Other ring networks have two counter-rotating rings that help them recover from network faults.
- FDDI is a fast fiber-optic network based on the ring topology

Advantages of Ring topology

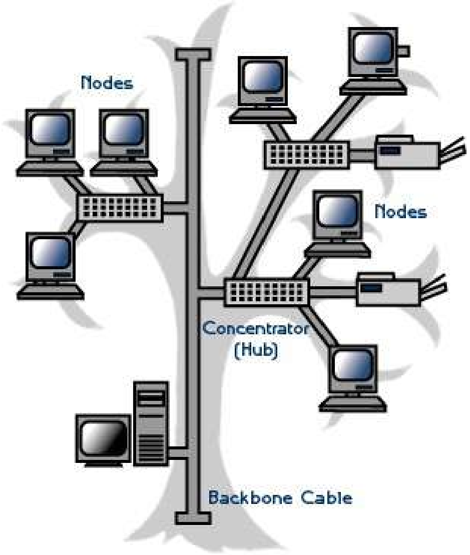
- The ring topology offers the following advantages:
  - Because every computer is given equal access to the token, no one computer can monopolize the network.
  - The fair sharing of the network allows the network to degrade gracefully (continue to function in a useful, if slower, manner rather than fail once capacity is exceeded) as more users are added.

Disadvantages of Ring topology

- The ring topology has the following disadvantages:
  - Failure of one computer on the ring can affect the whole network.
  - It is difficult to troubleshoot a ring network.
  - Adding or removing computers disrupts the network.

Tree topology

- A tree topology combines characteristics of linear bus and star topologies. It consists of groups of star-configured workstations connected to a linear bus backbone cable (See fig).
- Tree topologies allow for the expansion of an existing network, and enable schools to configure a network to meet their needs.



Advantages of a Tree Topology

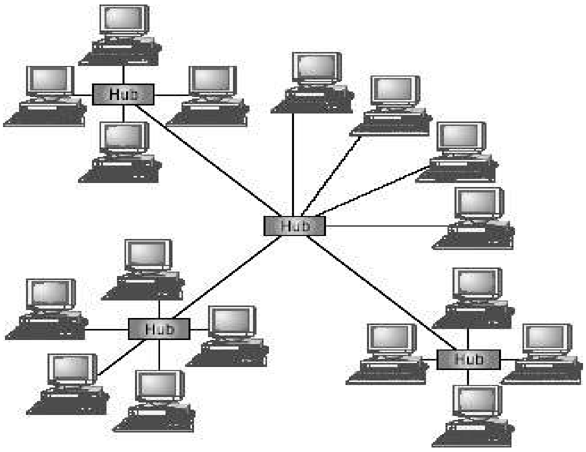
- Point-to-point wiring for individual segments.
- Supported by several hardware and software vender.

Disadvantages of a Tree Topology

- Overall length of each segment is limited by the type of cabling used.
- If the backbone line breaks, the entire segment goes down.
- More difficult to configure and wire than other topologies.

Hybrid topology

- You can expand a star network by placing another star hub where a computer might otherwise go, allowing several more computers or hubs to be connected to that hub. This creates a *hybrid star* network, like the one shown.



Advantages

- It is easy to modify and add new computers to a star network without disturbing the rest of the network. You simply run a new line from the computer to the central location and plug it into the hub. When the capacity of the central hub is exceeded, you can replace it with one that has a larger number of ports to plug lines into.
- You can use several cable types in the same network with a hub that can accommodate multiple cable types.
- The center of a star network is a good place to diagnose network faults.
- Intelligent hubs (hubs with microprocessors that implement features in addition to repeating network signals) also provide for centralized monitoring and management of the network.
- Single computer failures do not necessarily bring down the whole star network. The hub can detect a network fault and isolate the offending computer or network cable and allow the rest of the network to continue operating.

Disadvantages

- If the central hub fails, the whole network fails to operate.
- Many star networks require a device at the central point to rebroadcast or switch network traffic.
- It costs more to cable a star network because all network cables must be pulled to one central point, requiring more cable than other networking topologies.

6.3 MEDIA ACCESS CONTROL

What is Media Access Control?

A media access control is a network data transfer policy that determines how data is transmitted between two computers. The media access control policy involves sub-layers of the data link layer 2 in the OSI reference model.

The medium access control (MAC) is a sublayer of the data link layer of the OSI reference model for data transmission. It is responsible for flow control and multiplexing for transmission medium. It controls the transmission of data packets via remotely shared channels.

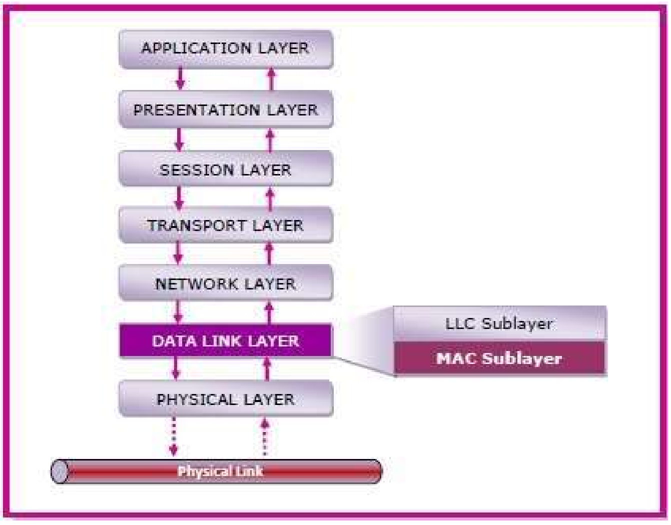
The essence of the MAC protocol is to ensure non-collision. A collision takes place when two or more terminals transmit data/information simultaneously. This leads to a breakdown of communication, which can prove costly for organizations that lean heavily on data transmission.

MAC Layer in the OSI Model

The Open System Interconnections (OSI) model is a layered networking framework that conceptualizes how communications should be done between heterogeneous systems. The data link layer is the second lowest layer. It is divided into two sublayers –

- The logical link control (LLC) sublayer
- The medium access control (MAC) sublayer

The following diagram depicts the position of the MAC layer –



Functions of MAC Layer

- It provides an abstraction of the physical layer to the LLC and upper layers of the OSI network.
- It is responsible for encapsulating frames so that they are suitable for transmission via the physical medium.
- It resolves the addressing of source station as well as the destination station, or groups of destination stations.
- It performs multiple access resolutions when more than one data frame is to be transmitted. It determines the channel access methods for transmission.
- It also performs collision resolution and initiating retransmission in case of collisions.
- It generates the frame check sequences and thus contributes to protection against transmission errors.

MAC Addresses

MAC address or media access control address is a unique identifier allotted to a network interface controller (NIC) of a device. It is used as a network address for data transmission within a network segment like Ethernet, Wi-Fi, and Bluetooth.

MAC address is assigned to a network adapter at the time of manufacturing. It is hardwired or hard-coded in the network interface card (NIC). A MAC address comprises of six groups of two hexadecimal digits, separated by hyphens, colons, or no separators. An example of a MAC address is 00:0A:89:5B:F0:11.

Media Access Control Methods

This network channel through which data is transmitted between terminal nodes to avoid collision has three various ways of accomplishing this purpose. They include:

- Carrier sense multiple access with collision avoidance (CSMA/CA)
- Carrier sense multiple access with collision detection (CSMA/CD)
- Demand priority
- Token passing

Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA)

Carrier sense multiple access with collision avoidance (CSMA/CA) is a media access control policy that regulates how data packets are transmitted between two computer nodes. This method avoids collision by configuring each computer terminal to make a signal before transmission. The signal is carried out by the transmitting computer to avoid a collision.

Multiple access implies that many computers are attempting to transmit data. Collision avoidance means that when a computer node transmitting data states its intention, the other waits at a specific length of time before resending the data.

CSMA/CA is data traffic regulation is slow and adds cost in having each computer node signal its intention before transmitting data. It used only on Apple networks.

Carrier Sense Multiple Access with Collision Detection (CSMA/CD)

Carrier sense multiple access with collision detection (CSMA/CD) is the opposite of CSMA/CA. Instead of detecting data to transmit signal intention to prevent a collision, it observes the cable to detect the signal before transmitting.

Collision detection means that when a collision is detected by the media access control policy, transmitting by the network stations stops at a random length of time before transmitting starts again.

It is faster than CSMA/CA as it functions in a network station that involves fewer data frames being transmitted. CSMA/CD is not as efficient as CSMA/CA in preventing network collisions. This is because it only detects huge data traffic in the network cable. Huge data traffic increases the possibility of a collision taking place. It is used on the Ethernet network.

Demand Priority

The demand priority is an improved version of the Carrier sense multiple access with collision detection (CSMA/CD). This data control policy uses an 'active hub' in regulating how a network is accessed. Demand priority requires that the network terminals obtain authorization from the active hub before data can be transmitted.

Another distinct feature of this MAC control policy is that data can be transmitted between the two network terminals at the same time without collision. In the Ethernet media, demand priority directs that data is transmitted directly to the receiving network terminal.

Token Passing

This media access control method uses free token passing to prevent a collision. Only a computer that possesses a free token, which is a small data frame, is authorized to transmit. Transmission occurs from a network terminal that has a higher priority than one with a low priority.

Token passing flourishes in an environment where a large number of short data frames are transmitted. This media access control policy is highly efficient in avoiding a collision. Possession of the free token is the only key to transmitting data by a network node. Each terminal holds this free token for a specific amount of time if the network with the high priority does not have data to transmit, the token is passed to the adjoining station in the network.

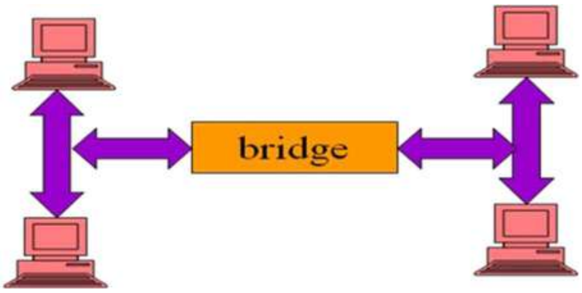
Media access control regulates how a network is accessed by computer terminals and transmits from one terminal to the other without collision. This is achieved through CSMA/CD, CSMA/CA, demand priority, or Token passing.

6.4 Use of Repeaters, Bridges, Router, Gateways, Public Network, X.25, Frame Relay

Use of Repeaters

- A network device used to regenerate or replicate a signal. Repeaters are used in transmission systems to regenerate analog or digital signals distorted by transmission loss.
- Analog repeaters frequently can only amplify the signal while digital repeaters can reconstruct a signal to near its original quality.
- In a data network, a repeater can relay messages between sub networks that use different protocols or cable types.
- Hubs can operate as repeaters by relaying messages to all connected computers. A repeater cannot do the intelligent routing performed by bridges and routers.
- Network **repeaters** regenerate incoming electrical, wireless or optical signals. With physical media like Ethernet or Wi-Fi, data transmissions can only span a limited distance before the quality of the signal degrades.
- Repeaters attempt to preserve signal integrity and extend the distance over which data can safely travel.
- A repeater connects two segments of your network cable.
- It retimes and regenerates the signals to proper amplitudes and sends them to the other segments.
- When talking about, Ethernet topology, you are probably talking about using a hub as a repeater.
- Repeaters require a small amount of time to regenerate the signal. This can cause a propagation delay which can affect network communication when there are several repeaters in a row.
- Many network architectures limit the number of repeaters that can be used in a row.
- Repeaters work only at the physical layer of the OSI network model.
- Actual network devices that serve as repeaters usually have some other name.
- **Active hubs**, for example, are repeaters. Active hubs are sometimes also called "multiport repeaters," but more commonly they are just "hubs."
- Other types of "passive hubs" are not repeaters. In Wi-Fi, access points function as repeaters only when operating in so-called "repeater mode."

Bridge



- ➢ Bridge is physical device typically a box of two port which connect two network at Data Link Layer.
- ➢ A Bridge provides packet filtering at Data link layer .
- ➢ A bridge to join to existing LAN or two split one LAN in to two segment. Bridge operate in promiscuous mode.
- ➢ Data packet enter the bridge through either one of the port and the bridge then read the destination address in each packet header and decides how to process that packet . This is called packet filtering.
- ➢ If the destination address of a packet arriving from one network segment is that of a computer on the other segment, the bridge Tx it out from other port.
- ➢ If the destination address of a computer on a same network segment as the computer that generated it, the bridge discard the packet.

Bridges &Collision

- ➢ A collisions domain is a network that is constructed so that when two computers transmit packet at the same time a collision occurs. When we add the new hub in existing network that the same collision domain as the original network because Hub relay the signal without filtering the packet
- ➢ Bridge do not relay the signals to other network until they have receive the entire packet. For this reason two computer on different side of the bridge do not cause to conflict.
- ➢ Bridge maintain the internal address table that listed the hardware address of the computer on both segment . When bridge receive the packet and read the destination address DLL header. It check the address against its lists. If the address is associated with the segment other than that from which the packet arrived, the bridge relay it to that segment there are two type of bridge

  (a) Local Bridge

  (b)Translation Bridge

  (c) Remote Bridge

Local Bridge

- ➢ Standard type of bridge use to connect network segment of same type and same location is called local bridge. This is simplest type of bridge ,it does not modify the data in packet. It simply read the address in the data link layer protocol and pass the packet or discard it.

Translation bridge

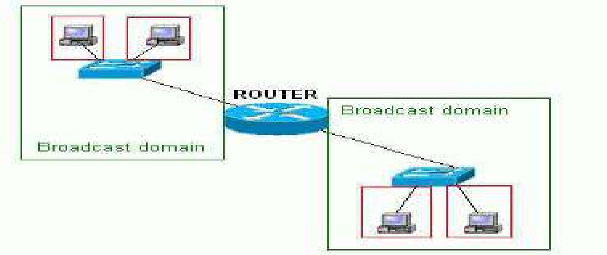
- It is DLL device that connect network using different network media or different protocol. This bridge is more complicated than local bridge. The bridge can thus connect an Ethernet segment to FDDI (Fiber Distributed Data Interface) segment or connect two different type of Ethernet type segment such as (100BaseTx).

Remote Bridge

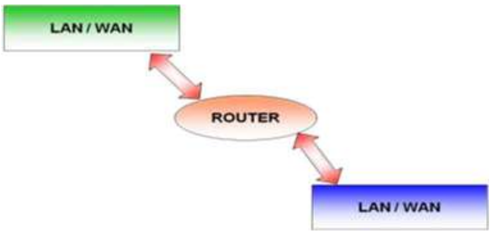
- Remote bridge is designed to connect two network segment at distance locations using some form of wide area network link. The link can be a modem connection leased telephone line or any type of WAN technology. The advantage of using a bridge in this manner is that you reduce the amount of traffic passing over the WAN link., which is usually far slower and more expensive than the local Network.

Router

- Routers are packet forwarding devices or it is a device that forwards data packet along network.
- Routers are located at gateway the places where two or more networks connect. .

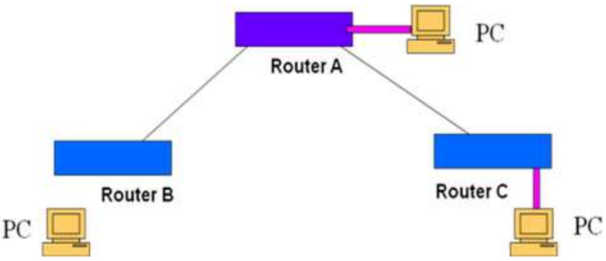


- A router is connected to at least two networks, commonly two LANs or WAN or a LAN and its ISP's network. Routers allow transmission of data between network segments.
- Routers are specialized computers that send your messages and those of every other Internet user speeding to their destinations along thousands of pathways.



Function Of Router

- Routing is the process of moving data throughout a network , passing through several network segments.
- Router gets information about which path to take from files on the routers called routing tables. These table contain information about which router network interface to place information on in order to send it to a particular network segment. Routers will not pass unknown or broadcast packets. A router will route a packet only if it has a specific destination.



- Keeping the messages moving
- Transmitting packets
- Knowing where to send data
- Understanding the protocols
- Tracing message

The main difference between routers, bridges

- A router passes packets by looking up the destination in it's routing table of an incoming packet. Bridges work at layer 1/2 physical media where everything is passed from one port to another with no regard for source, destination, or network address.
- Routers work at layer 3 moving packets from one port to another based on the L3 address - i.e. IP address, IPX address, etc.

6.5 CARRIER SENSE MULTIPLE ACCESS(CSMA)

- It was developed to minimize the chance of collision and to increase the performance.
- It requires that each station first listen to the medium before sending.
- It is based on the principle sense before transmitor listen before talk.
- It can reduce the possibility of collision, but it cannot eliminate it.
- In CSMA a station senses the carrier on the channel before starting its own transmission.
- The vulnerable time for CSMA is the propagation time Tp.
- propagation time needed for a signal to propagate from one end of the medium to the other. When a station sends a frame, and any other station tries to send a frame during this time, a collision will result. When a channel is sense to be idle, a station can take one of the three different approaches to transmit a packet on to the channel. These three protocols are as follows:

Non-persistent CSMA

- In non-persistent CSMA, when a station having a packet to transmit and finds that the channel is busy, it backs off for a fixed interval of time.
- It then checks the channel again and if the channel is free then it transmits.
- The back-off delay is determined by the transmission of a frame, propagation time and other system parameter.

- If the channel is already in use, the station does not continuously sense it for the purpose of seizing it immediately upon detecting the end of the previous transmission. But it waits a random period of time and again check for activity.
-

1-Persistent CSMA

- Any station wishing to transmit, monitor the channel continuously until the channel is idle and then transmit with probability one, hence the name 1-persistent.
- When two or more stations are waiting to transmit, a collision is guaranteed. Since each station will transmit immediately at the end of busy period. In this case each will wait a random amount of time and will then reattempt to transmit.
- As in the case with non-persistent CSMA, the performance of 1-persistent CSMA protocol depends only on the channel delay time.

P-Persistent CSMA:-

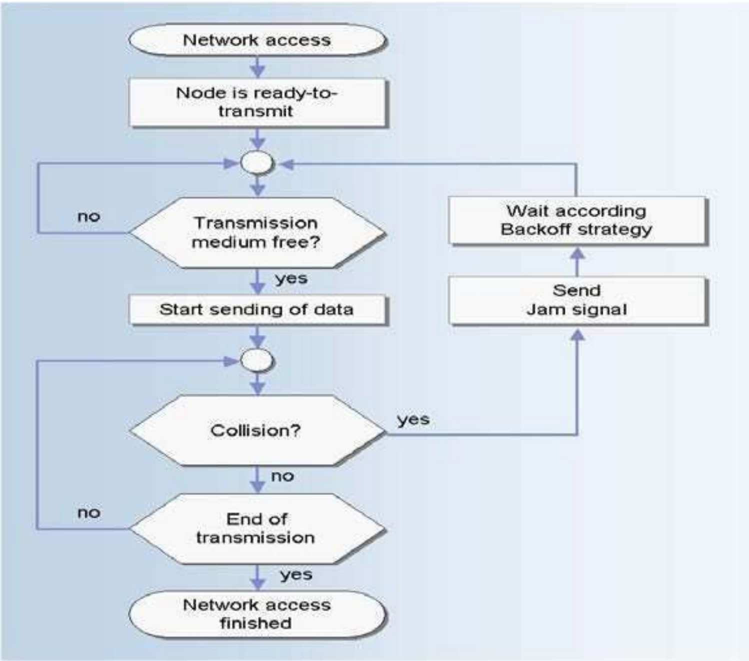
- To reduce the probability of collision in 1-persistent CSMA, not all allowed transmitting immediately, after the channel is idle.
- When a station becomes ready to send and its sense the channel to be idle, it either transmits with a probability p or it defers transmission by one time slot with a probability q=1-p.if the differed slot is also idle, the station either transmits with probability p or defers again with a probability q.this process is repeated until either packets are transmitted of the channel is busy.

CSMA with collision detection (CSMA/CD)

- CSMA/CD is the most commonly used protocol for LANs. CSMA/CD specifications were developed jointly by digital equipment corporation (DEC), Intel, and Xerox. This network is called as Ethernet. The IEEE802.3
- CSMA/CD stands for LAN are based on Ethernet specification.
- The basic protocol is that, a station with a message to send must monitor the channel to see if any other station is sending. If another station is sending, the second station must wait or defer, until the sending station has finished. Then it may send its message. If no station was sending at

the time that it first listened, the station may send its message immediately. The term carrier sense indicates this listening before transmitting behaviour.

- **Carrier Sense Multiple Access/Collision Detect** (CSMA/CD) is the protocol for carrier transmission access in Ethernet networks.
- **Carrier-sense multiple access with collision detection** describes how the Ethernet protocol regulates communication among nodes
- On Ethernet, any station can send a frame at any time. Each station senses whether the medium is idle and therefore available for use. If it is, the station begins to transmit its first frame. If another station also tries to transmit at the same time, a collision occurs and the frames are discarded and then a jamming signal is sent throughout the network in order to notify all stations of the collision.
- Each station then waits for a random period of time and retries. If another collision occurs, the time intervals from which the random waiting time is selected are increased step by step. This is known as exponential back off. The stations retry until successful transmission of the frame.CSMA/CD is specified in the IEEE 802.3 standard.



- The **jam signal** is a signal that carries a 32-bit binary pattern sent by a data station to inform the other stations that they must not transmit.

- **Transmit –** The stations (nodes) sends the frames to other stations (nodes)
- **Carrier Sense –** The stations (nodes) listen to the medium if it is idle for transmission
- **Back off –** After collision occurs, a jam signal is sent to notify all stations of the collision. After the jam signal is sent, the stations (nodes) wait for a random period of time called Back off period.
- If two or more stations have message to send at the same time and they are separated by significant distances on the bus/channel. each may begin transmitting at roughly the same time without being aware of the other station .the signal from each station will superimpose on the channel and is garbled beyond the decoding ability of the receiving station .this is termed as collision .
- A protocol is required for transmitting station to monitor the channel while sending each of its messages and to detect such collisions.
- When a collision has been detected, each of sending stations must cease transmitting, wait for a random length of time, and then try again. Because of quick termination of transmission time and bandwidth is saved. Therefore CSMA/CD is more efficient than ALOHA and CSMA. CSMA/CD network work best on a bus, multipoint topology with busty asynchronous transmission.CSMA/CD has totally decentralized control and is based on connection access.

CSMA with collision avoidance (CSMA/CA)

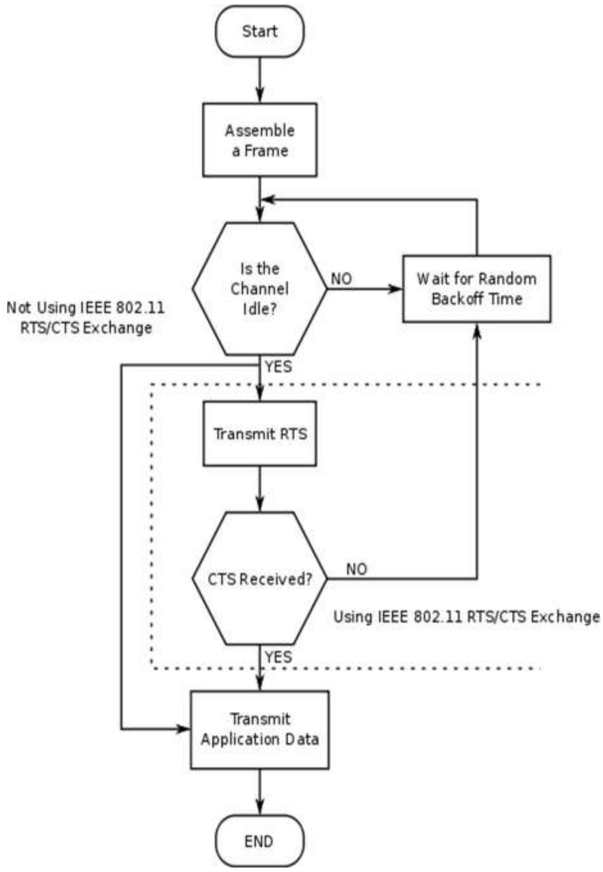
Collision avoidance is used to improve the performance of the CSMA method by attempting to divide the channel somewhat equally among all transmitting nodes within the collision domain. **1. Carrier Sense**

- Prior to transmitting, a node first listens to the shared medium (such as listening for wireless signals in a wireless network) to determine whether another node is transmitting or not.

2. Collision Avoidance

- If another node was heard, we wait for a period of time for the node to stop transmitting before listening again for a free communications channel.

- Request to Send/Clear to Send (RTS/CTS) may optionally be used at this point to mediate access to the shared medium. This goes some way to alleviating the problem of hidden nodes because, for instance, in a wireless network, the Access Point only issues a Clear to send to one node at a time.
- Transmission: if the medium was identified as being clear or the node received CTS to explicitly indicate it can send, it sends the frame in its entirety.
- Unlike CSMA/CD, it is very challenging for a wireless node to listen at the same time as it transmits (its transmission will dwarf any attempt to listen).
- Although CSMA/CA has been used in a variety of wired communication systems, it is particularly beneficial in a wireless LAN due to a common problem of multiple stations being able to see the Access Point, but not each other. This is due to differences in transmitting power, and receives sensitivity.



Start
Assemble a Frame
Is the Channel Idle?
NO
Wait for Random Backoff Time
Not Using IEEE 802.11 RTS/CTS Exchange
YES
Transmit RTS
CTS Received?
NO
Using IEEE 802.11 RTS/CTS Exchange
YES
Transmit Application Data
END

- CSMA/CA performance is based largely upon the modulation technique used to transmit the data between nodes. Studies show that under ideal propagation conditions (simulations), Direct Sequence Spread Spectrum (DSS) provides the highest throughput for all nodes on a network when used in conjunction with CSMA/CA and the IEEE 802.11 RTS/CTS exchange under light network load conditions.

6.6 INTRODUCTION TO WIRELESS LAN

Wireless LAN stands for **Wireless Local Area Network**. It is also called LAWN (**Local Area Wireless Network**).

Wireless LANs (WLANs) are wireless computer networks that use high-frequency radio waves instead of cables for connecting the devices within a limited area forming LAN (Local Area Network). Users connected by wireless LANs can move around within this limited area such as home, school, campus, office building, railway platform, etc.

The IEEE 802.11 group of standards defines the technologies for wireless LANs. For path sharing, 802.11 standard uses the Ethernet protocol and CSMA/CA (carrier sense multiple access with collision avoidance). It also uses an encryption method i.e. wired equivalent privacy algorithm.

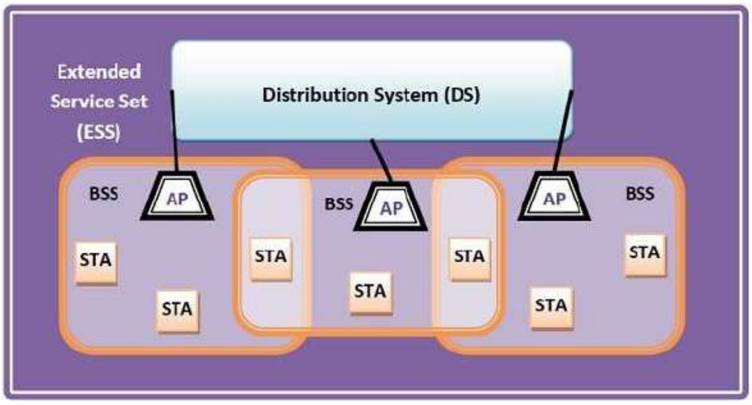
Wireless LANs provide high speed data communication in small areas such as building or an office. WLANs allow users to move around in a confined area while they are still connected to the network.

In some instance wireless LAN technology is used to save costs and avoid laying cable, while in other cases, it is the only option for providing high-speed internet access to the public. Whatever the reason, wireless solutions are popping up everywhere.

Components of WLANs

The components of WLAN architecture as laid down in IEEE 802.11 are −

- **Stations (STA)** − Stations comprises of all devices and equipment that are connected to the wireless LAN. Each station has a wireless network interface controller. A station can be of two types −
  - Wireless Access Point (WAP or AP)
  - Client
- **Basic Service Set (BSS)** − A basic service set is a group of stations communicating at the physical layer level. BSS can be of two categories −
  - Infrastructure BSS
  - Independent BSS
- **Extended Service Set (ESS)** − It is a set of all connected BSS.
- **Distribution System (DS)** − It connects access points in ESS.



Types of WLANS

WLANs, as standardized by IEEE 802.11, operates in two basic modes, infrastructure, and ad hoc mode.

- **Infrastructure Mode** − Mobile devices or clients connect to an access point (AP) that in turn connects via a bridge to the LAN or Internet. The client transmits frames to other clients via the AP.
- **Ad Hoc Mode** − Clients transmit frames directly to each other in a peer-to-peer fashion.

Advantages of WLANs

- They provide clutter-free homes, offices and other networked places.
- The LANs are scalable in nature, i.e. devices may be added or removed from the network at greater ease than wired LANs.
- The system is portable within the network coverage. Access to the network is not bounded by the length of the cables.

  - **Flexibility:** Within radio coverage, nodes can communicate without further restriction. Radio waves can penetrate walls, senders and receivers can be placed anywhere (also non-visible, e.g., within devices, in walls etc.).
  - **Planning:** Only wireless ad-hoc networks allow for communication without previous planning, any wired network needs wiring plans.
  - **Design:** Wireless networks allow for the design of independent, small devices which can for example be put into a pocket. Cables not only restrict users but also designers of small notepads, PDAs, etc.
  - **Robustness:** Wireless networks can handle disasters, e.g., earthquakes, flood etc. whereas, networks requiring a wired infrastructure will usually break down completely in disasters.
  - **Cost:** The cost of installing and maintaining a wireless LAN is on average lower than the cost of installing and maintaining a traditional wired LAN, for two reasons. First, after providing wireless access to the wireless network via an access point for the first user, adding additional users to a network will not increase the cost. And second, wireless LAN eliminates the direct costs of cabling and the labor associated with installing and repairing it.
  - **Ease of Use:** Wireless LAN is easy to use and the users need very little new information to take advantage of WLANs.

Disadvantages of WLANs

- Since radio waves are used for communications, the signals are noisier with more interference from nearby systems.
- Greater care is needed for encrypting information. Also, they are more prone to errors. So, they require greater bandwidth than the wired LANs.
- WLANs are slower than wired LANs.

- **Quality of Services:** Quality of wireless LAN is typically lower than wired networks. The main reason for this is the lower bandwidth due to limitations is radio transmission, higher error rates due to interference and higher delay/delay variation due to extensive error correction and detection mechanisms.
- **Proprietary Solutions:** Due to slow standardization procedures, many companies have come up with proprietary solutions offering standardization functionality plus many enhanced features. Most components today adhere to the basic standards IEEE 802.11a or 802.11b.
- **Restrictions:** Several govt. and non-govt. institutions world-wide regulate the operation and restrict frequencies to minimize interference.
- **Global operation:** Wireless LAN products are sold in all countries so, national and international frequency regulations have to be considered.
- **Low Power:** Devices communicating via a wireless LAN are typically power consuming, also wireless devices running on battery power. Whereas the LAN design should take this into account and implement special power saving modes and power management functions.
- **License free operation:** LAN operators don't want to apply for a special license to be able to use the product. The equipment must operate in a license free band, such as the 2.4 GHz ISM band.
- **Robust transmission technology:** If wireless LAN uses radio transmission, many other electrical devices can interfere with them (such as vacuum cleaner, train engines, hair dryers, etc.).Wireless LAN transceivers cannot be adjusted for perfect transmission is a standard office or production environment.

Unit-6

LAN TECHNOLOGY

1. Topology and Transmission Media
2. Medium Access control
3. Bridges, Hub, Switch
4. Ethernet (CSMA/CD), Fiber Channel
5. Wireless LAN Technology

6.1
Topology and Transmission Media

Topology

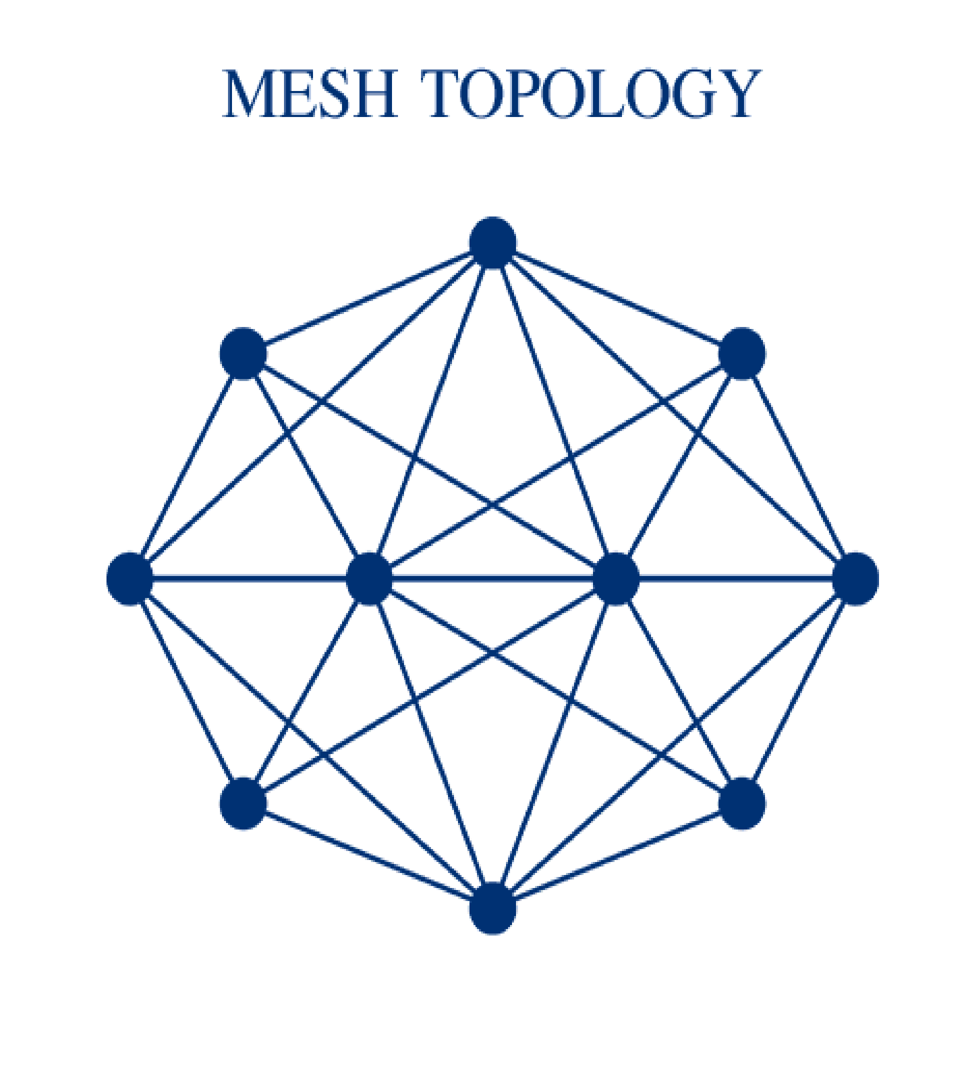
- The way in which the connections are made of the physical devices is called the topology of the network.
- Network topology specifically refers to the physical layout of the network, especially the locations of the computers and how the cable is run between them.
- It is important to select the right topology for how the network will be used.
- Each topology has its own strengths and weaknesses.
- The four most common topologies are
  - Mesh topology
  - Bus topology
  - Star topology
  - Ring topology
  - Tree Topology

MESH Topology

A network setup where each computer and network device is interconnected with one another, allowing for most transmissions to be distributed, even if one of the connections go down.

This topology is not commonly used for most computer networks as it is difficult and expensive to have redundant connection to every computer.

However, this topology is commonly used for wireless networks



MESH Topology

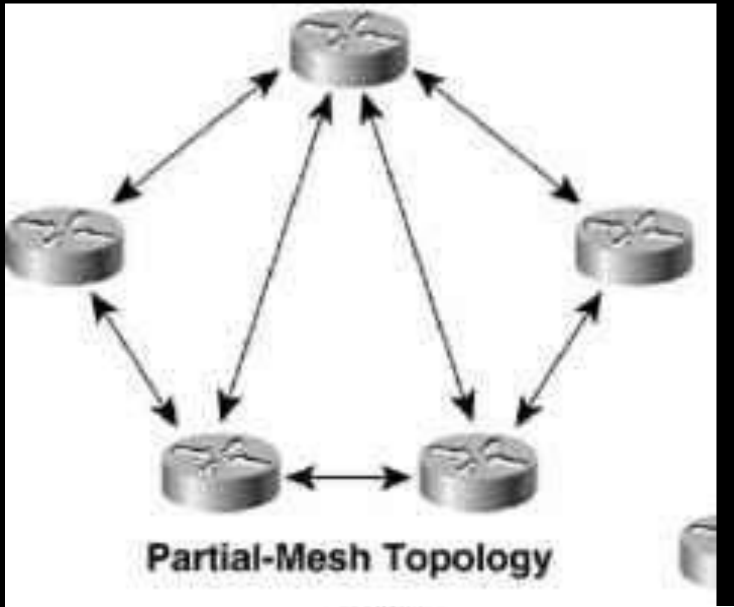
In MESH Topology, there exists multiple paths between hosts.

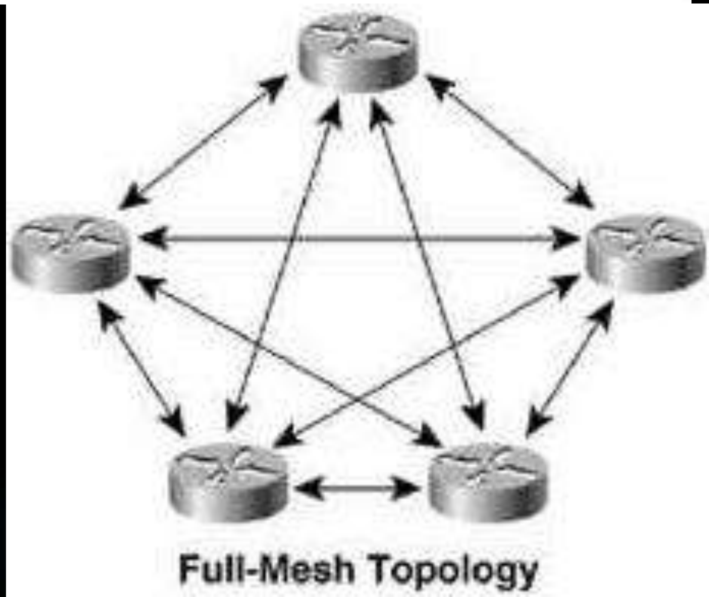
Some hosts are directly connected to some other nodes.

There exist multiple paths between the sender and the receiver.

This mesh topology can be connected in two forms i.e.
(i) partially
(ii) fully.

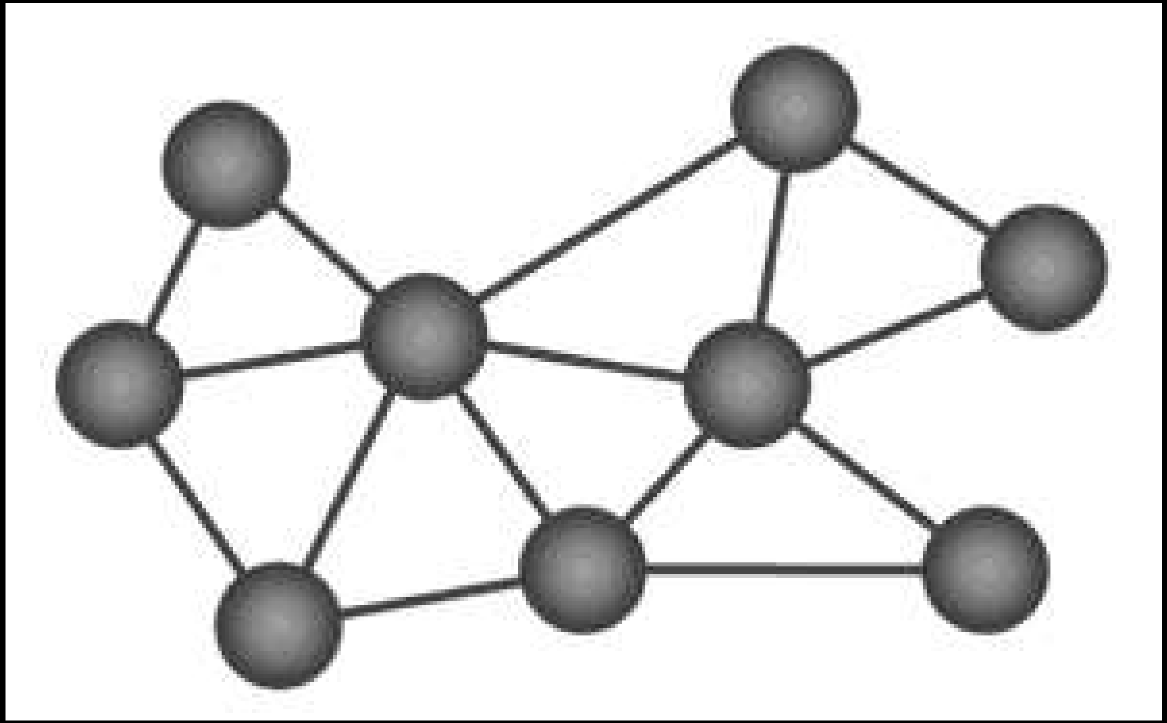
In this mesh topology, there is no break in communication





Advantages

- Reliable and no traffic problem
- Privacy and Security is maintained
- No failure of data transmission
- Data can be transmitted from different devices simultaneously.
- This topology can withstand high traffic.
- If one of the components fails there is always an alternative present.



Disadvantages

- Require extensive cabling
- High cost
- High chances of redundancy in many of the network connections.
- Set-up and maintenance of this topology is very difficult.
- Even administration of the network is tough

BUS Topology

In BUS topology, all data transmissions are carried out by a single common cable.

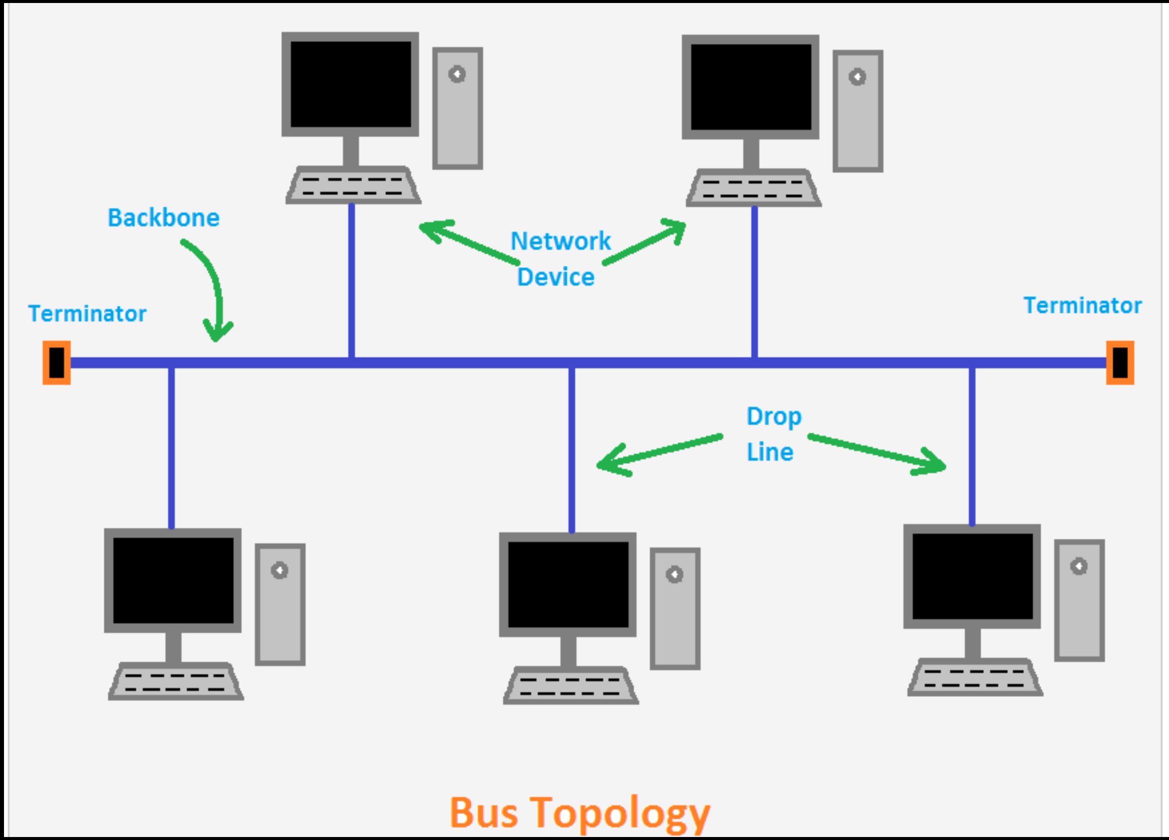
The single common cable is known as BUS.

All communication from one computer to another is done by the common cable or BUS.

Due to this, only one computer can transmit data at a given time.

If two computers try to send data at the same time, then collision occurs and both the computers retract back and try after sometime.

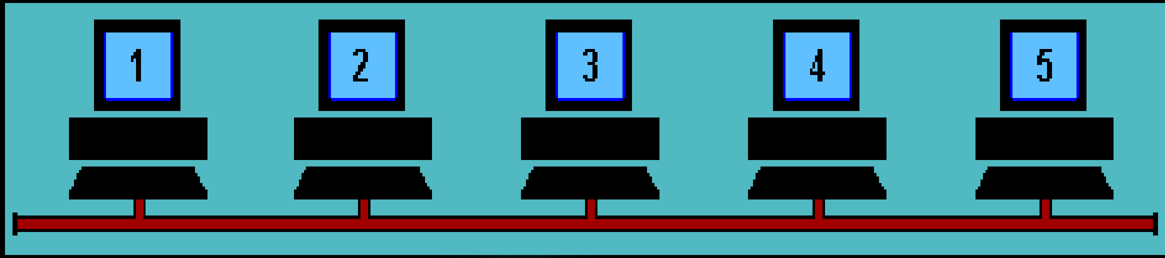
There are rules or protocols that determine which computer has right to transmit.



Each communicating device is connected to the BUS with the help of special line, called dropline.

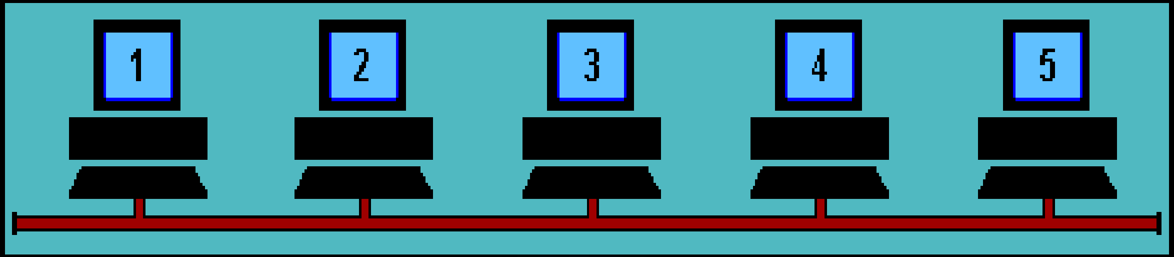
BUS Topology

- The bus topology is often used when a network installation is small, simple, or temporary.
- On a typical bus network, the cable is just one or more wires, with no active electronics to amplify the signal or pass it along from computer to computer.
- This makes the bus a passive topology.
- When one computer sends a signal up (and down) the wire, all the computers on the network receive the information, but only one (the one with the address that matches the one encoded in the message) accepts the information. The rest disregard the message.
- Only one computer at a time can send a message; therefore, the number of computers attached to a bus network can significantly affect the speed of the network.
- A computer must wait until the bus is free before it can transmit



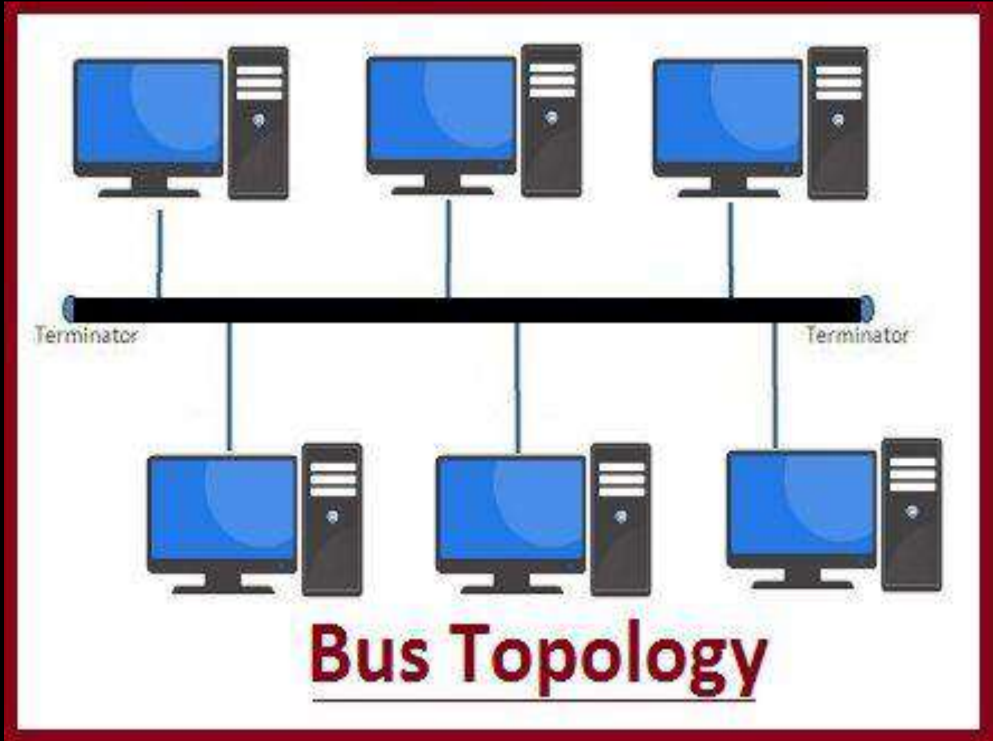
BUS Topology

- Another important issue in bus networks is termination. Since the bus is a passive topology, the electrical signal from a transmitting computer is free to travel the entire length of the cable.
- Without termination, when the signal reaches the end of the wire, it bounces back and travels back up the wire.
- When a signal echoes back and forth along an unterminated bus, it is called ringing.
- To stop the signals from ringing, you attach terminators at either end of the segment.
- The terminators absorb the electrical energy and stop the reflections.
- Cables cannot be left unterminated in a bus network.
- *Ethernet 10Base2 (also known as thinnet) is an inexpensive network based on the bus topology*



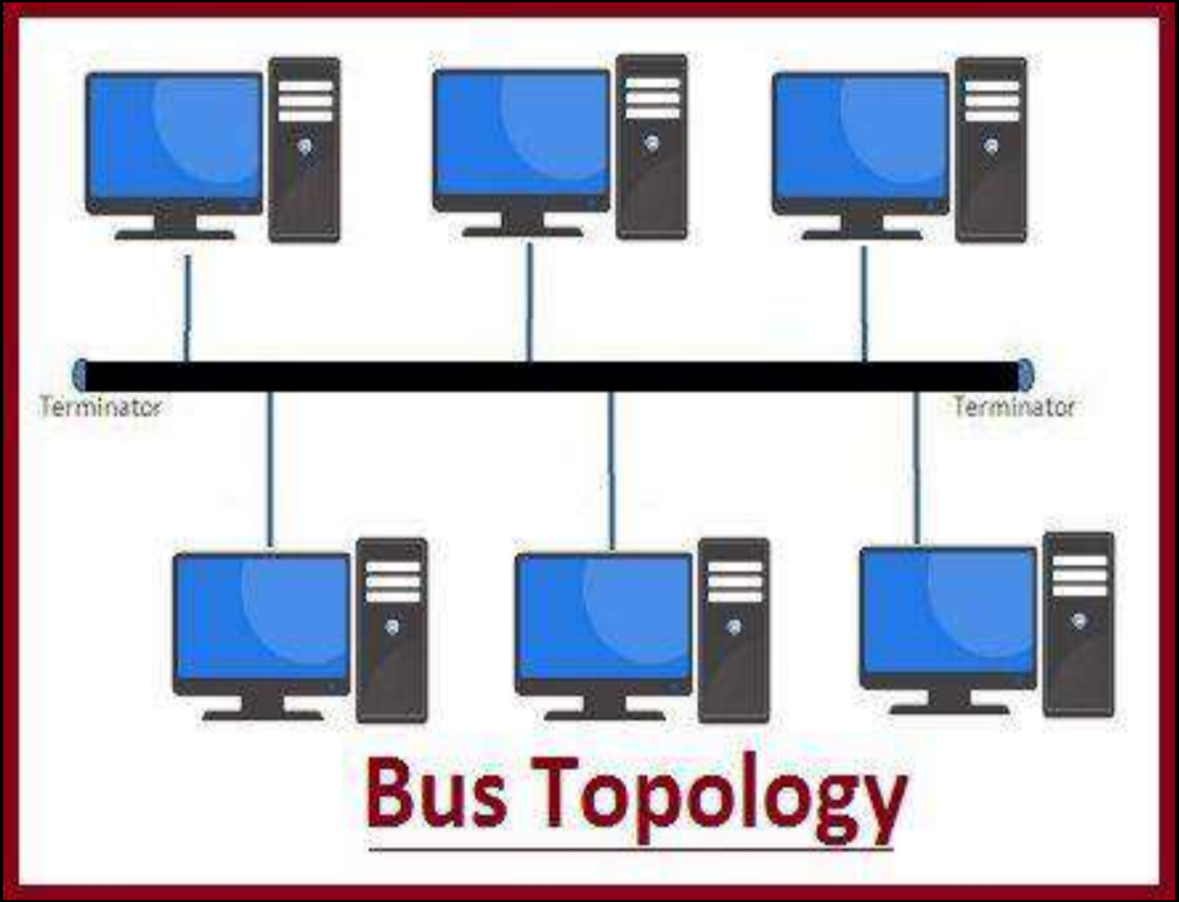
Advantages

i) Easy to add or remove computer without disturbing others
ii) One message can be sent at any point of time
iii) It uses broadcast system (n sender sends data to all other hosts of the network)
iv) Economical and easy to setup
v) Mainly Co-Axial Cable is used
vi) Used in small LANs
vii) Requires the least amount of cable
viii) It is easy to extend a bus using BNC connector.
ix) A repeater can also be used to extend; a repeater boosts the signal and allows it to travel a longer distance



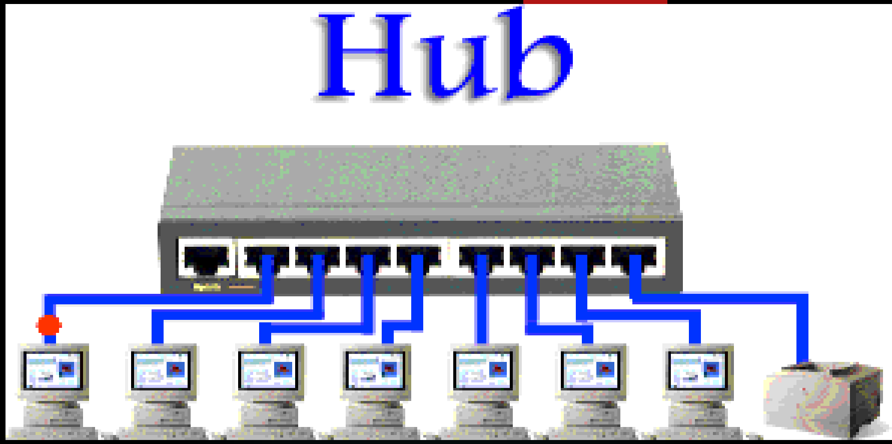
Disadvantages

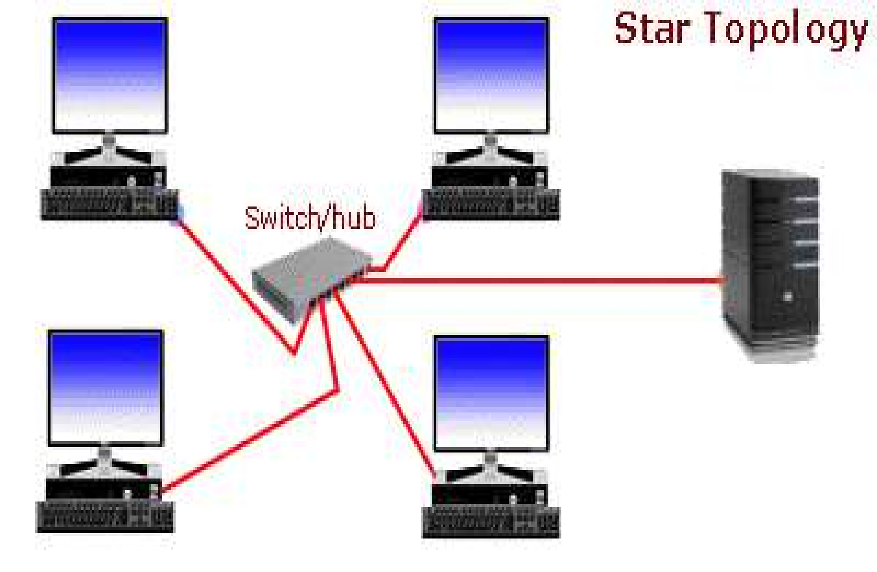
I. Heavy network traffic can slow a bus
II. a bus network with a lot of computers can spend a lot of its bandwidth
III. Each connector weakens the electrical signal, and too many may prevent the signal from being correctly received all along the bus.
IV. A cable break or malfunctioning computer anywhere between two computers can cause them not to be able to communicate with each other.
V. A cable break or loose connector will also cause reflections and bring down the whole network, causing all network activity to stop



STAR Topology

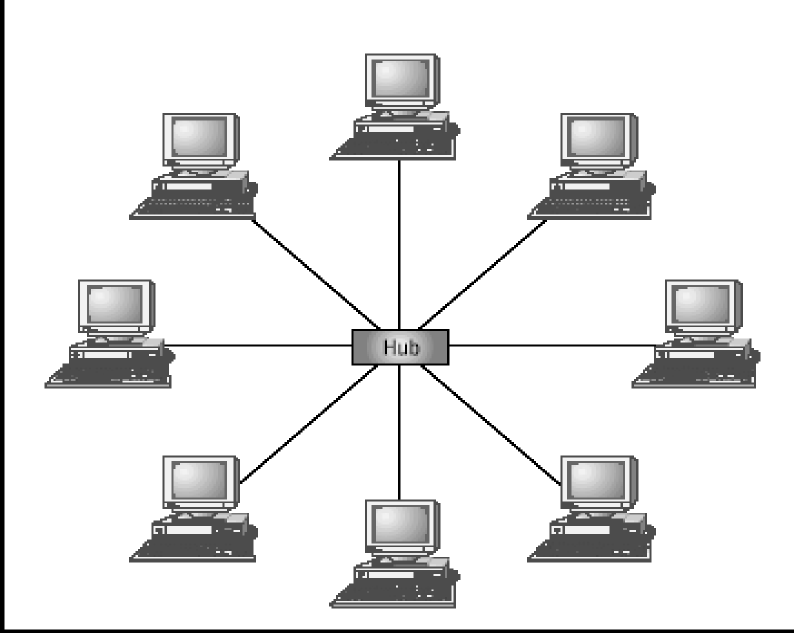
- Star Topology is the most popular topology that is used for setting up of LANs.
- In this topology, there is a central device known as Hub or Switch.
- Every host is connected to the central device.
- All communication goes through the central hub.
- Ex – Public telephone network
- In a star topology, all the cables run from the computers to a central location, where they are all connected by a device called a hub.
- Each computer on a star network communicates with a central hub that resends the message either to all the computers (in a broadcast star network) or only to the destination computer (in a switched star network).





STAR Topology

- The hub in a broadcast star network can be active or passive.
- An active hub regenerates the electrical signal and sends it to all the computers connected to it.
- This type of hub is often called a multiport repeater. Active hubs and switches require electrical power to run.
- A passive hub, such as wiring panels or punch-down blocks, merely acts as a connection point and does not amplify or regenerate the signal.
- Passive hubs do not require electrical power to run.
- Several types of cable can be used to implement a star network.
- In a star topology the computers are all connected by cables to a central point

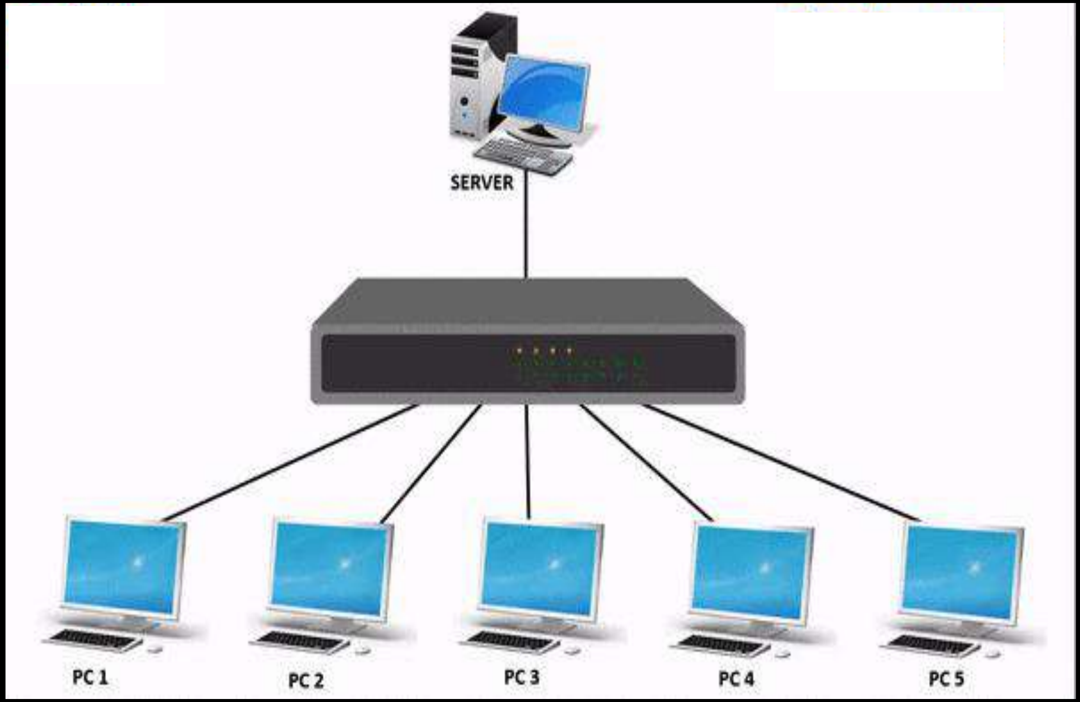


Characteristics

- It has a central device (hub / switch)
- Data transfers through the switch
- Each node is connected to the central device
- Twisted pair cable is used.

Advantages

- Less expensive
- Easy to install and reconfigure
- Nodes can be added or removed easily
- Data transmission is bidirectional
- If one link fails, then only that node is affected
- Easy fault identification
- High speed

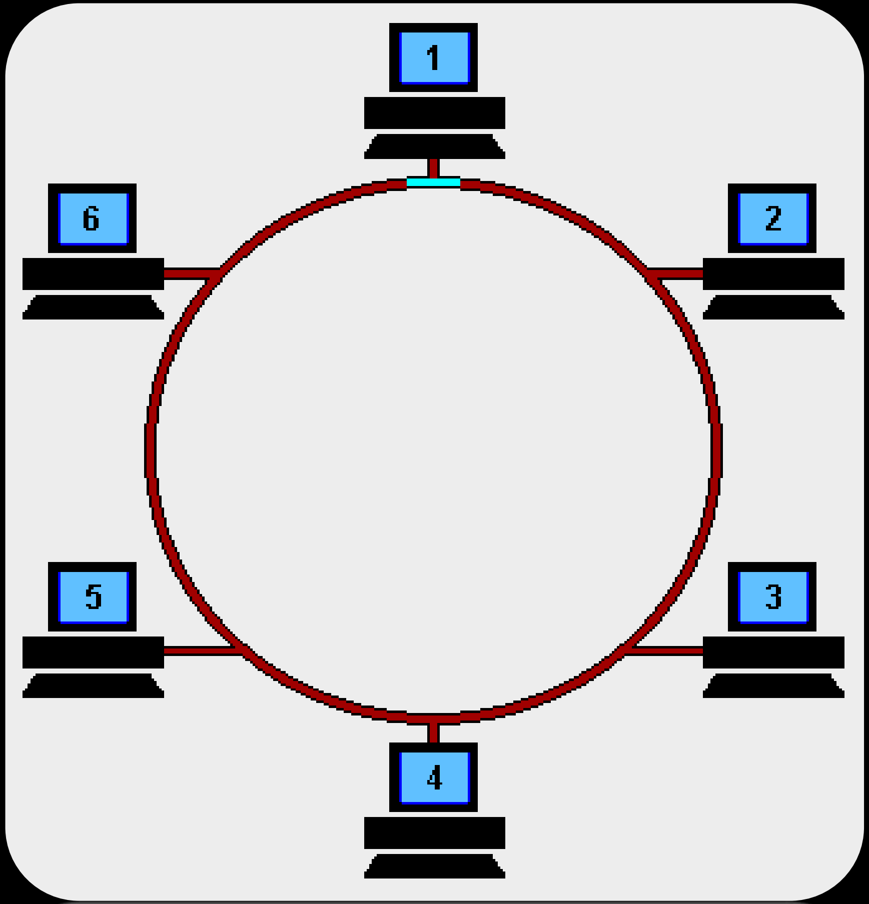


Disadvantages

- If the central hub is down, then the entire network goes down
- Requires more cabling than RING or BUS

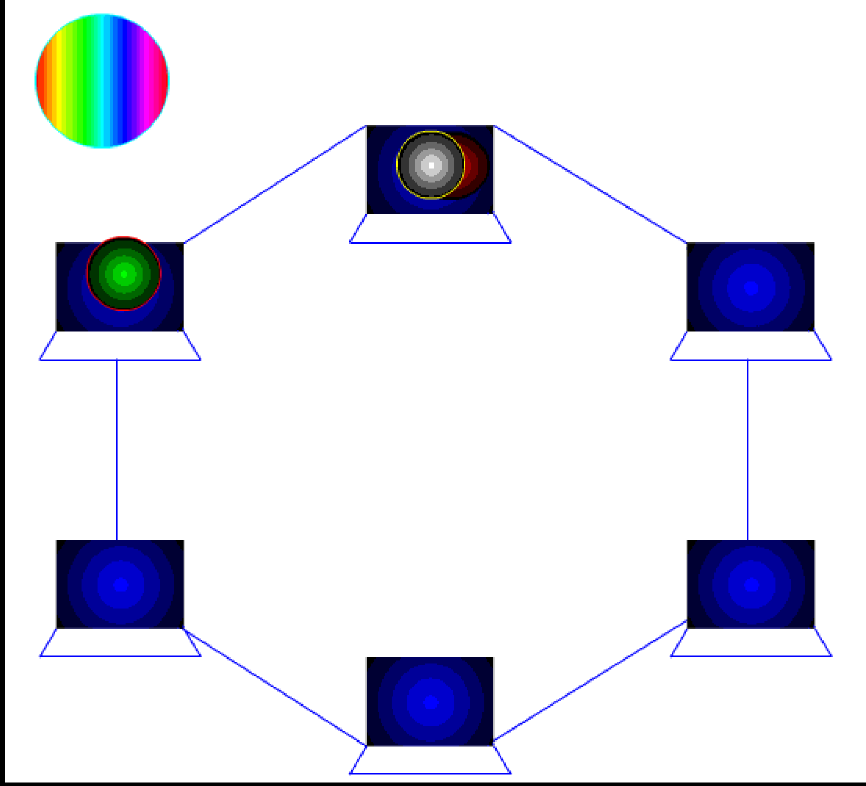
RING Topology

- In a Ring Topology, the communicating devices are connected with each other to form a circle.
- As they are in circle, every node has two neighbors.
- Every host is connected to its neighbors.
- In this topology, messages can travel only in one direction.
- It functions by passing packets from one node to another until it reached the destination.
- When a node receives a packet, it checks the destination address.
- If the address of the node matches with the address of the packet, then the node receives the packet.
- After receiving the packet, the node



RING Topology

- In a ring topology, each computer is connected to the next computer, with the last one connected to the first.
- Rings are used in high-performance networks, networks that bandwidth be reserved for time-sensitive features such as video and audio, or when even performance is needed when a large number of clients access the network.
- Every computer is connected to the next computer in the ring, and each retransmits what it receives from the previous computer.
- The messages flow around the ring in one direction.
- Since each computer retransmits what it receives, a ring is an *active* network and is not subject to the signal loss problems a bus experiences. There is no termination because there is no end to the ring.

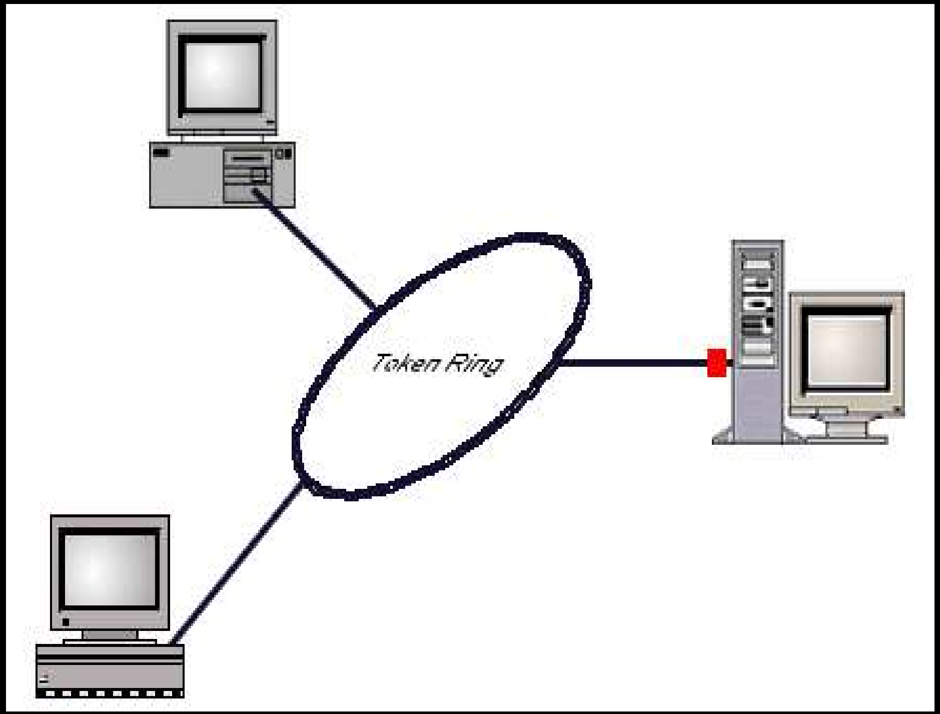


RING Topology

- Some ring networks do token passing. A short message called a token is passed around the ring until a computer wishes to send information to another computer.
- That computer modifies the token, adds an electronic address and data, and sends it around the ring.
- Each computer in sequence receives the token and the information and passes them to the next computer until either the electronic address matches the address of a computer or the token returns to its origin.
- The receiving computer returns a message to the originator indicating that the message has been received.
- The sending computer then creates another token and places it on the network, allowing another station to capture the token and begin transmitting.
- The token circulates until a station is ready to send and captures the token.
- This all happens very quickly: a token can circle a ring 200 meters in diameter at about 10,000 times a second. Some even faster networks circulate several tokens at once.
- Other ring networks have two counter-rotating rings that help them recover from network faults. FDDI is a fast fiber-optic network based on the ring topology

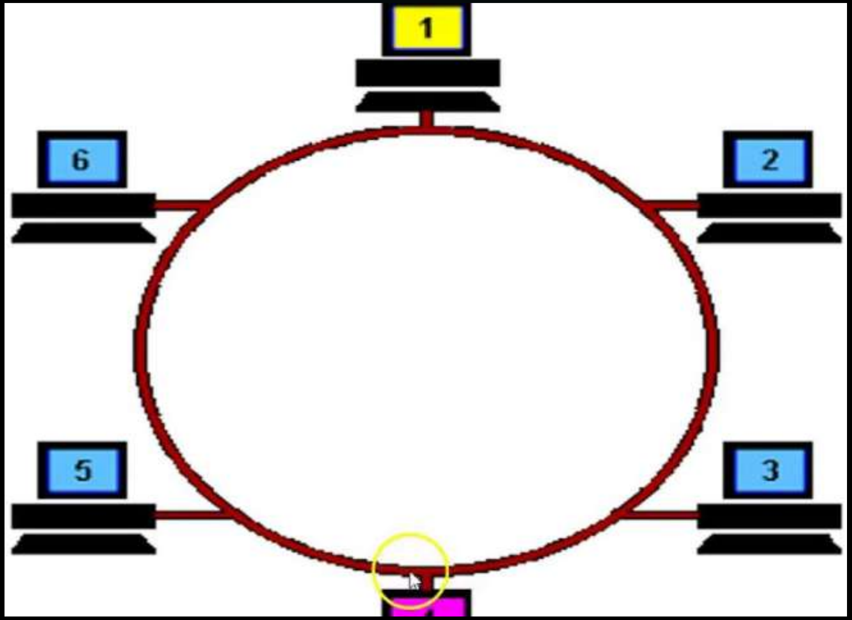
Characteristics

- Each node connected to the network via repeaters.
- Shared resources can be attached to any host
- Most rings have a monitor station, which has the task of removing corrupted / unwanted packets from the network.



Advantages

- Reliable.
- Each node is connected with two neighbors
- It is each to install and configure
- Because every computer is given equal access to the token, no one computer can monopolize the network.



Disadvantages

- Untraditional (data can transmit only in one direction)
- If cable between two hosts damaged, the entire network goes down.
- Failure of one computer on the ring can affect the whole network
- It is difficult to troubleshoot a ring network
- Adding or removing computers disrupts the network

TREE Topology

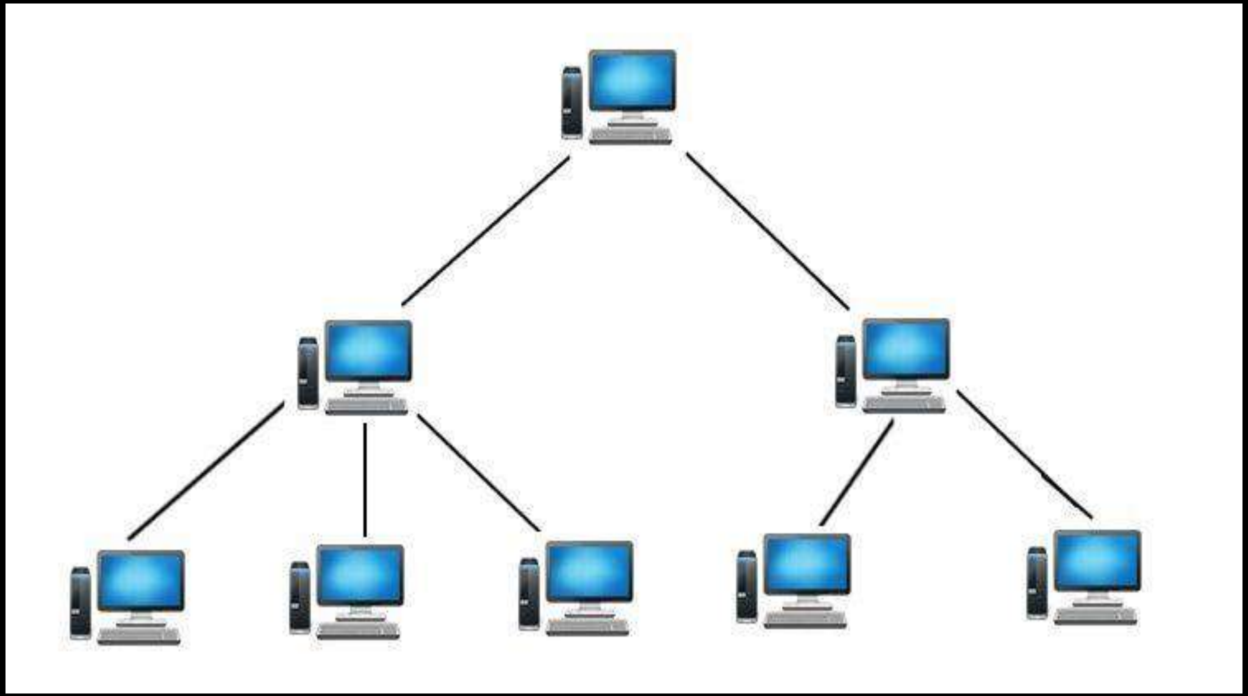
The TREE Topology joins different star network into a single network.

It also known as extended STAR topology.

It can be extended in such a manner that, instead of the node, a Hub can be connected to another Hub.

One Hub is the main hub whereas other are sub-hubs.

It gives the flexibility to increase the number of nodes in a network

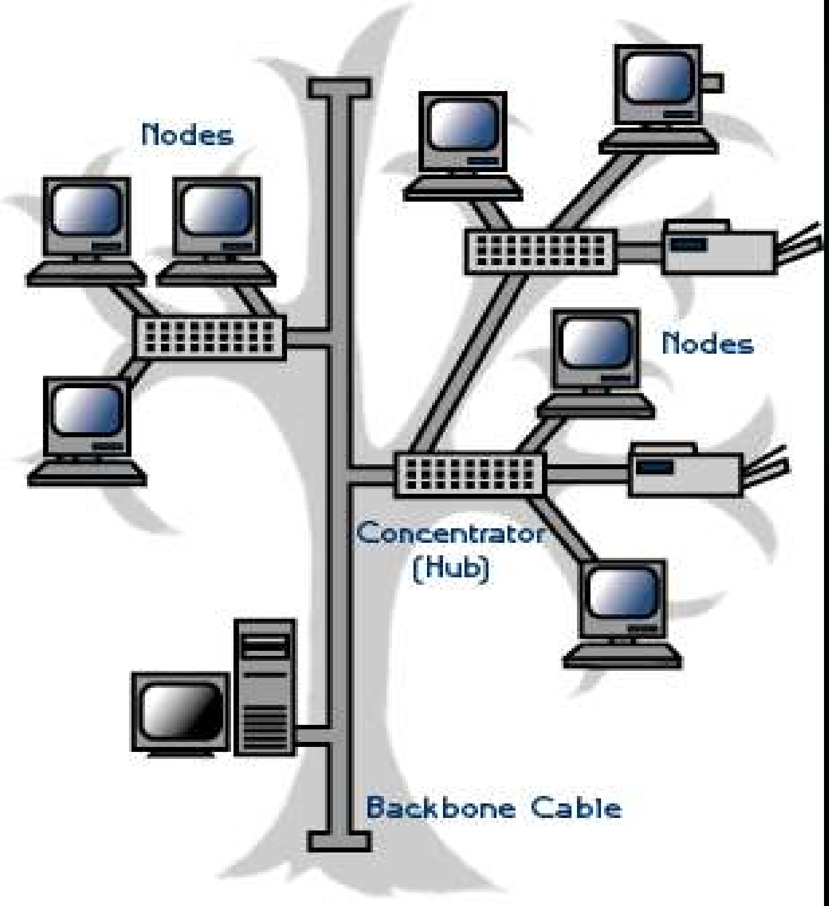


TREE Topology

A tree topology combines characteristics of linear bus and star topologies.

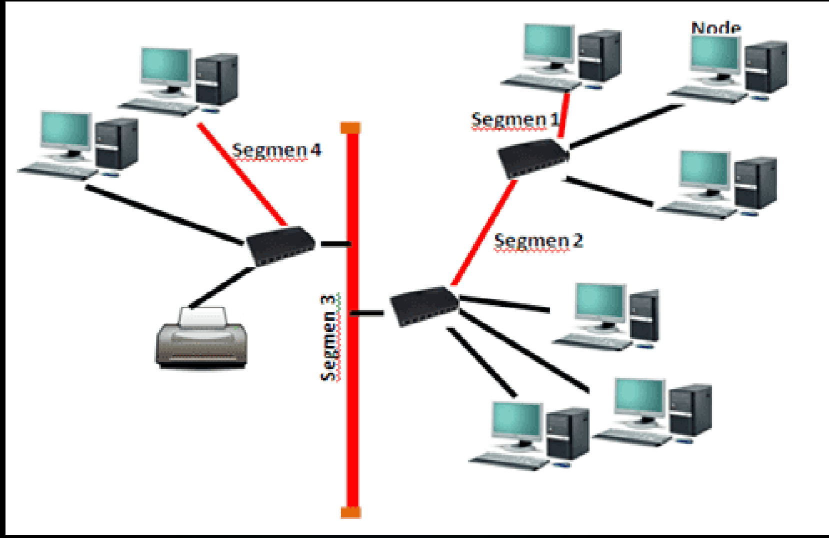
It consists of groups of star-configured workstations connected to a linear bus backbone cable.

Tree topologies allow for the expansion of an existing network, and enable schools to configure a network to meet their needs



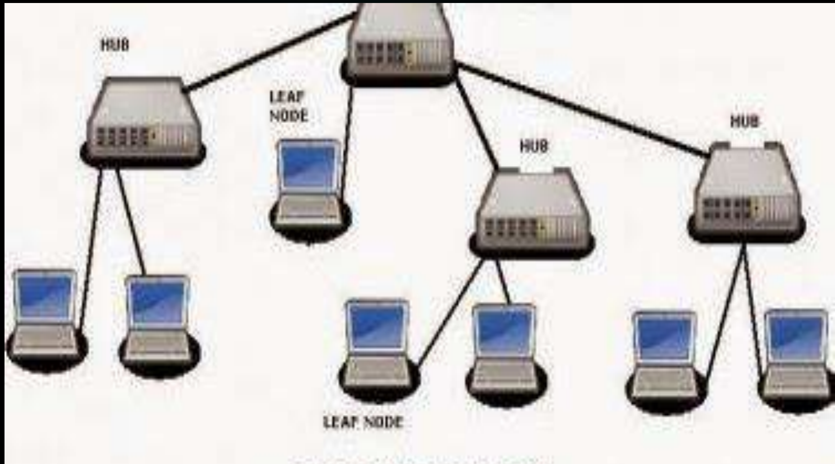
Advantages

- Different networks can be connected to form a single network
- Maintenance is easy
- Easy to add or remove at/from bottom of the tree
- Point-to-point wiring for individual segments.
- Supported by several hardware and software vender



Disadvantages

- If the top node is overloaded, then the network performance will become slow
- If the top is down, then the whole network will be down
- Overall length of each segment is limited by the type of cabling used
- More difficult to configure and wire than other topologies

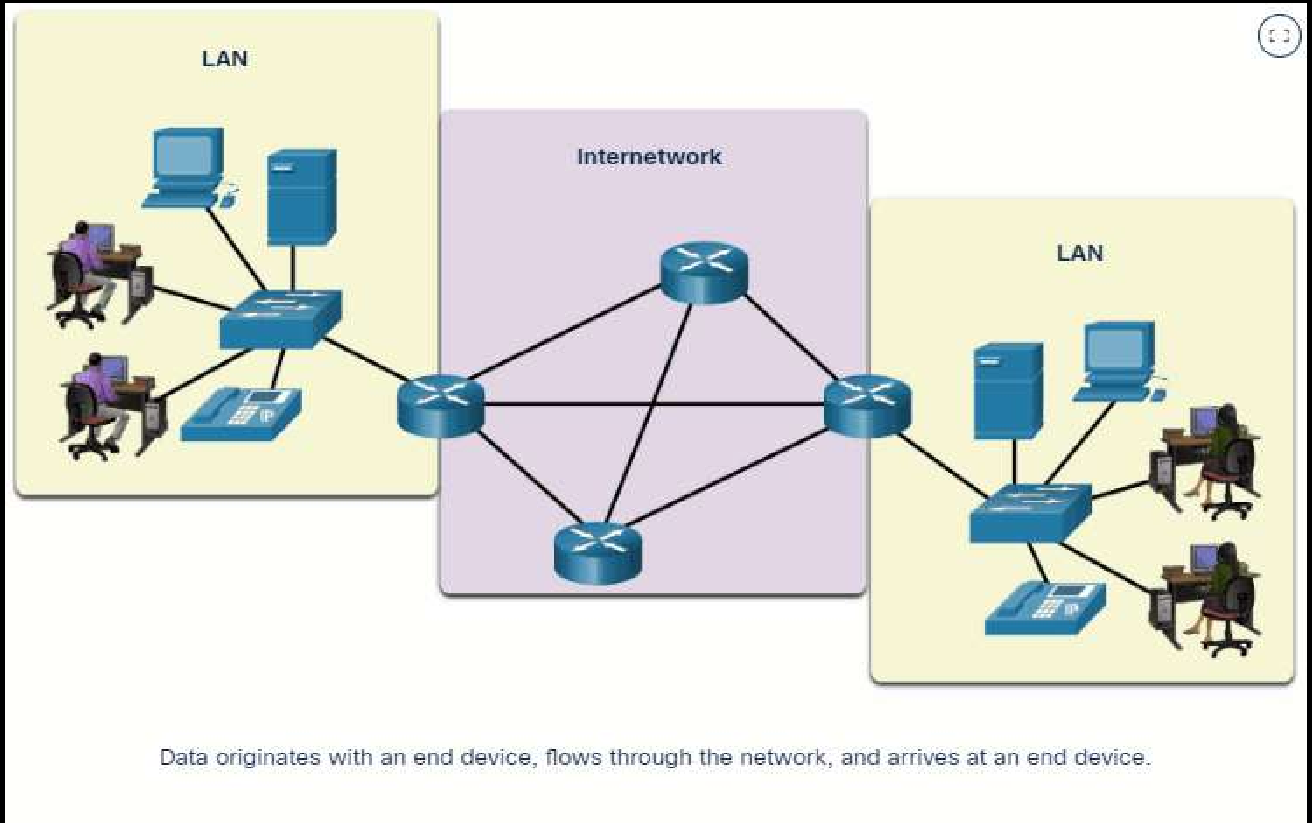


TREE TOPOLOGY

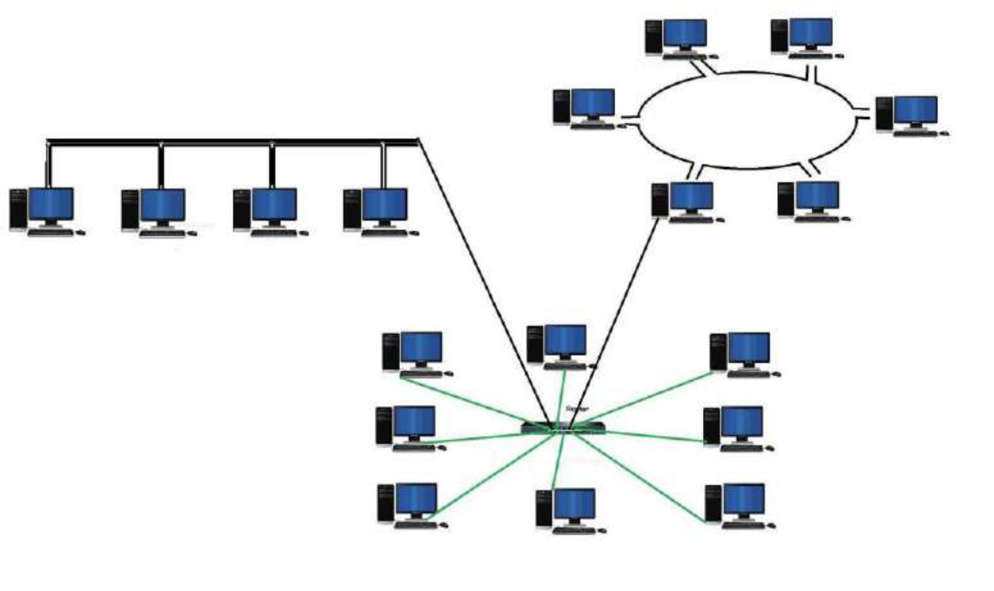
HYBRID Topology

A network that contains a combination of different topologies is considered as Hybrid topology.

The exact size, shape and the combination depends on the requirement



Data originates with an end device, flows through the network, and arrives at an end device.



Advantages

It is easy to modify and add new computers to a star network without disturbing the rest of the network.

If the capacity of the central hub is exceeded, it can be replaced with larger number of ports to plug lines.

Several cable types can be used in the same network with a hub that can accommodate multiple cable types.

The center of a star network is a good place to diagnose network faults.

Intelligent hubs (hubs with microprocessors that implement features in addition to repeating network signals) also provide for centralized monitoring and management of the network.

Single computer failures do not necessarily bring down the whole star network.
The hub can detect a network fault and isolate the offending computer or network cable and allow the rest of the network to continue operating

Disadvantages

If the central hub fails, the whole network fails to operate.

Many star networks require a device at the central point to rebroadcast or switch network traffic.

It costs more to cable a star network because all network cables must be pulled to one central point, requiring more cable than other networking topologies

6.3
Media Access Control

Media Access Control

What is Media Access Control?

- A media access control is a network data transfer policy that determines how data is transmitted between two computers.
- The media access control policy involves sub-layers of the data link layer 2 in the OSI reference model.
- The medium access control (MAC) is a sublayer of the data link layer of the OSI reference model for data transmission.
- It is responsible for flow control and multiplexing for transmission medium.

Media Access Control

What is Media Access Control?

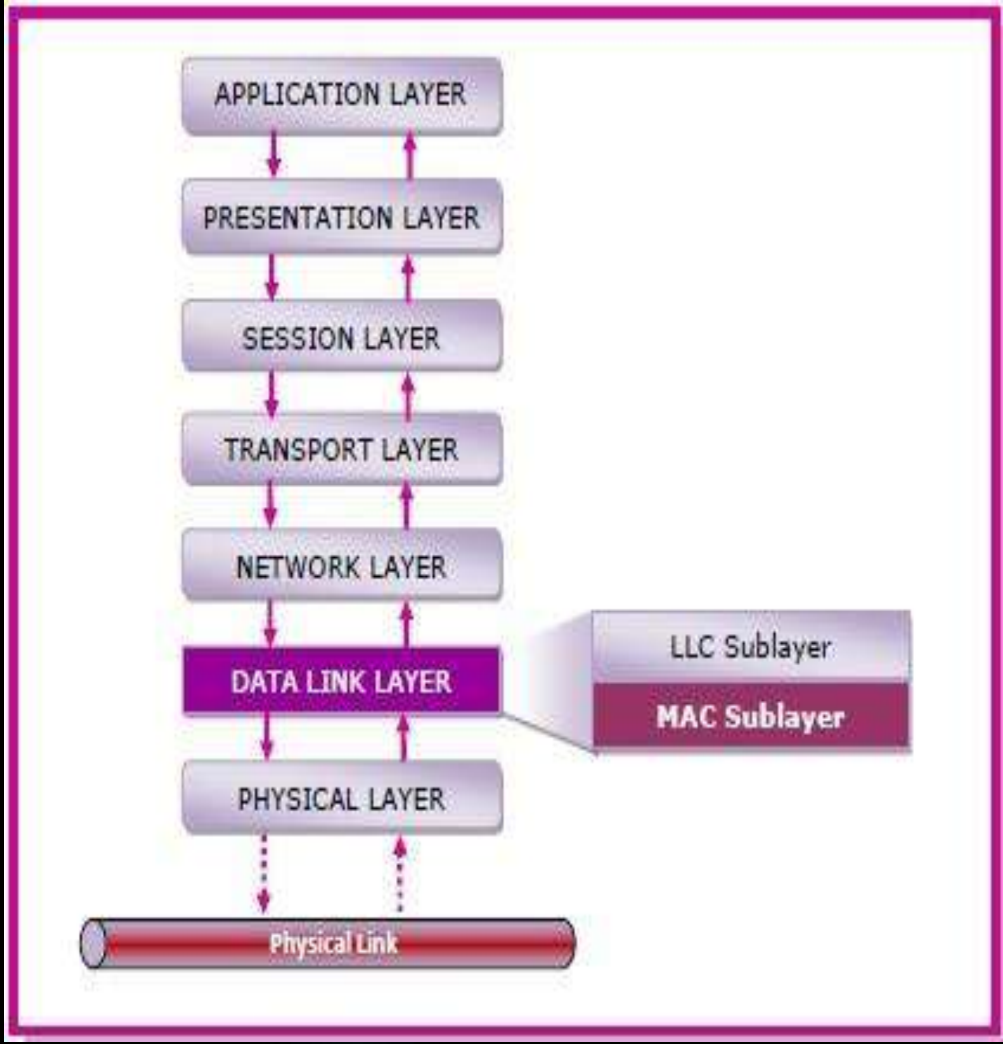
- It controls the transmission of data packets via remotely shared channels.
- The essence of the MAC protocol is to ensure non-collision.
- A collision takes place when two or more terminals transmit data/information simultaneously.
- This leads to a breakdown of communication, which can prove costly for organizations that lean heavily on data transmission

MAC Layer in OSI Model

Layer 2: Data Link
The data link layer is responsible for maintaining the data link between two hosts or nodes. Its characteristics and functions are as follows:

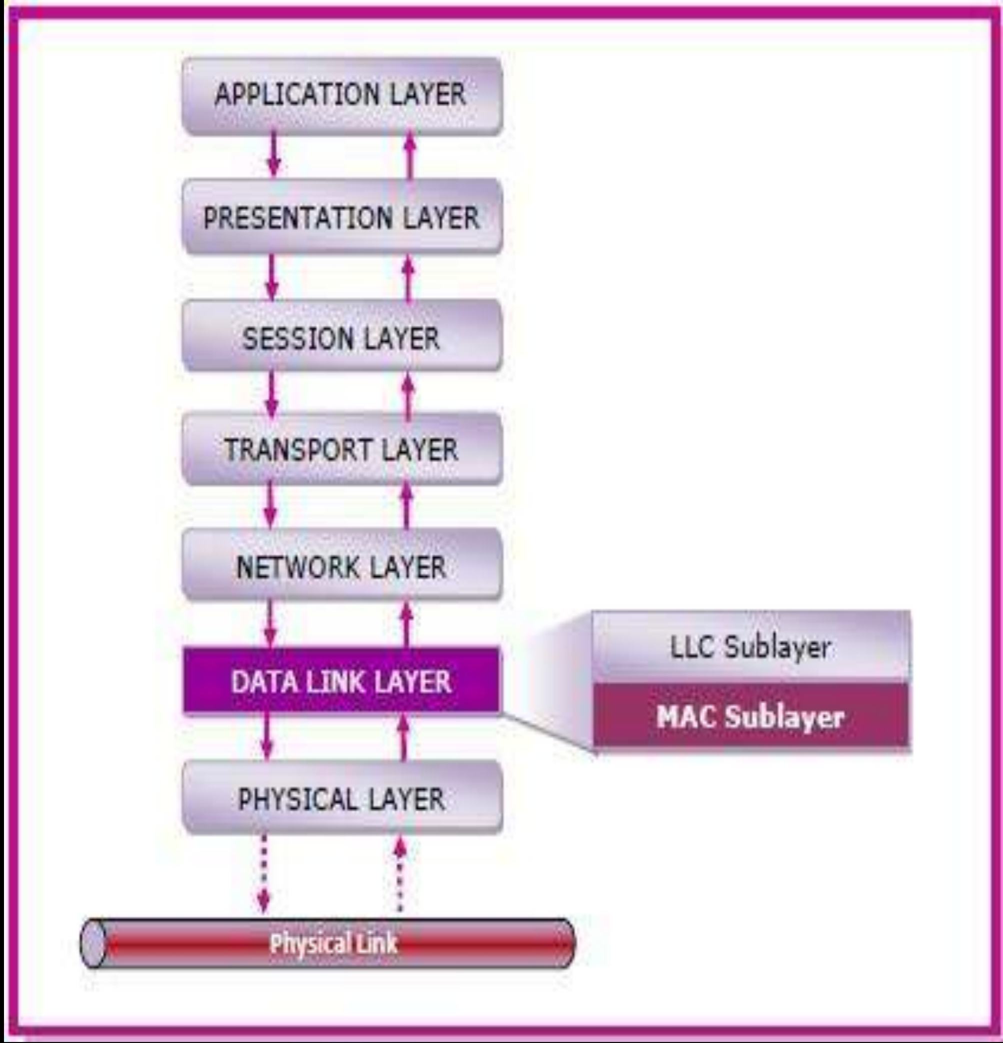
- Defines and manages the ordering of bits to and from data segments called packets
- Management of frames, which contains data arranged in an organized manner, which provides for an orderly and consistent method of sending data bits across the medium

▪



MAC Layer in OSI Model

- Responsible for flow control, which is the process of managing the timing of sending and receiving data so that it doesn't exceed the capacity of the physical connection
- Responsible for error notification, including receiving and managing error messaging related to physical delivery of packets
- Network devices that operate at this layer include Layer 2 switches (switching hubs) and bridges.
- It is divided into two sublayers
  - The logical link control (LLC) sublayer
  - The medium access control (MAC) sublayer



MAC Layer in OSI Model

- **Logical Link Control (LLC) sublayer :**

  The LLC sublayer provides the logic for the data link.
  Thus, it controls the synchronization, flow control, and error checking functions of the data link layer.

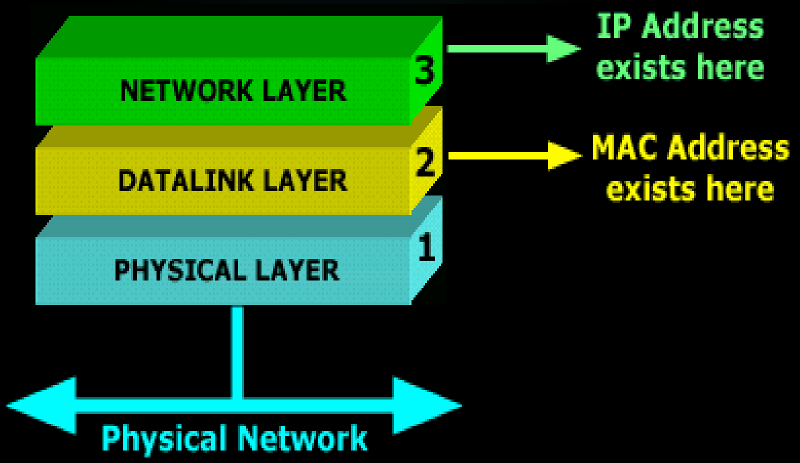
- **Media Access Control (MAC) sublayer :**

  The MAC sublayer provides control for accessing the transmission medium.
  It is responsible for moving data packets from one network interface card (NIC) to another, across a shared transmission medium.
  Physical addressing is handled at the MAC sublayer.
  MAC is also handled at this layer.
  This refers to the method used to allocate network access to computers and prevent them from transmitting at the same time, causing data collisions.
  Common MAC methods include Carrier Sense Multiple Access/Collision Detection (CSMA/CD), used by Ethernet networks, Carrier Sense Multiple Access/Collision Avoidance (CSMA/CA), used by AppleTalk networks, and token passing, used by Token Ring and Fiber Distributed Data Interface (FDDI) networks

Function of MAC Layer

Functions of MAC Layer

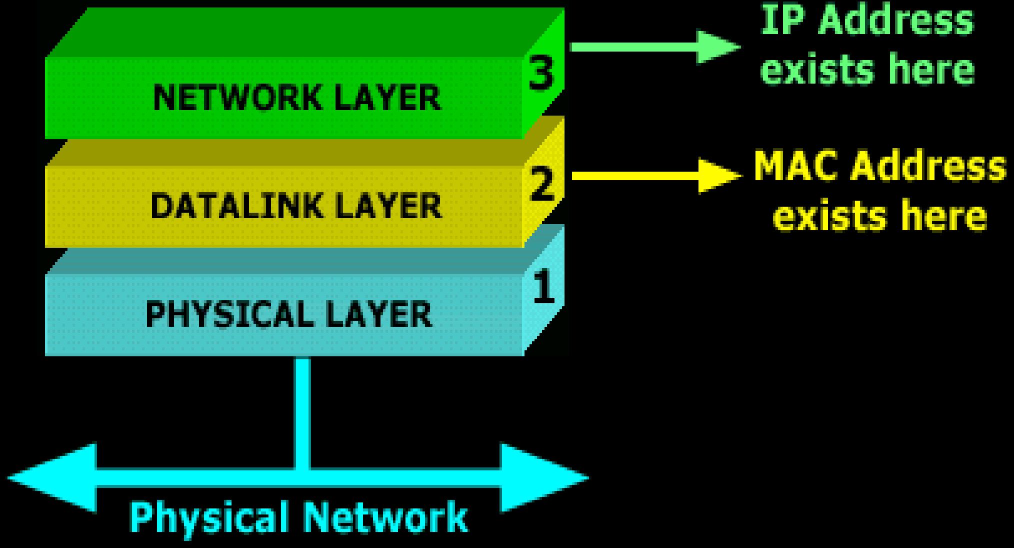
- It provides an abstraction of the physical layer to the LLC and upper layers of the OSI network.
- It is responsible for encapsulating frames so that they are suitable for transmission via the physical medium.
- It resolves the addressing of source station as well as the destination station, or groups of destination stations.
- It performs multiple access resolutions when more than one data frame is to be transmitted.
- It determines the channel access methods for transmission.
- It also performs collision resolution and initiating retransmission in case of collisions.
- It generates the frame check sequences and thus contributes to protection against transmission errors.



MAC Address

MAC Addresses

- MAC address or media access control address is a unique identifier allotted to a network interface controller (NIC) of a device.
- It is used as a network address for data transmission within a network segment like Ethernet, Wi-Fi, and Bluetooth.
- MAC address is assigned to a network adapter at the time of manufacturing.
- It is hardwired or hard-coded in the network interface card (NIC).
- A MAC address comprises of six groups of two hexadecimal digits, separated by hyphens, colons, or no separators.
- An example of a MAC address is 00:0A:89:5B:F0:11.



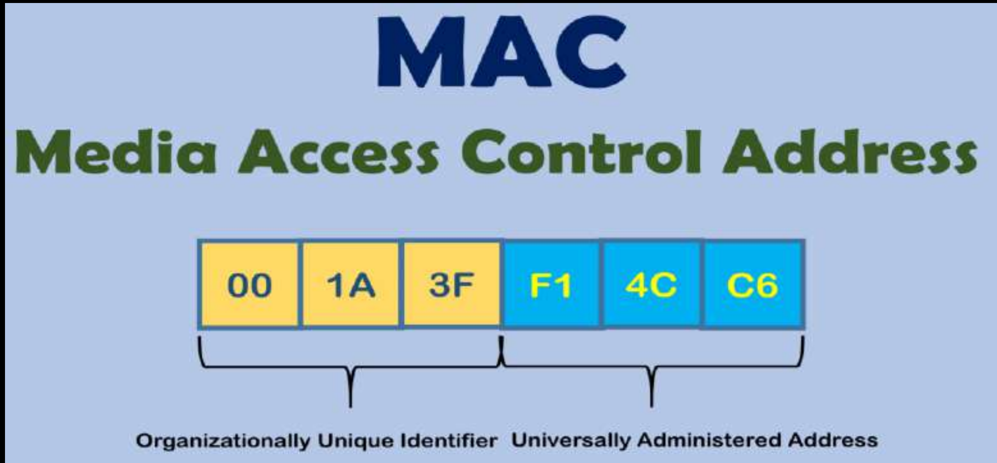
Media Access Control Methods

Media Access Control Methods

The network channel through which data is transmitted between terminal nodes to avoid collision has various ways of accomplishing this purpose.

They include:

- Carrier sense multiple access with collision avoidance **(CSMA/CA)**
- Carrier sense multiple access with collision detection **(CSMA/CD)**
- **Demand priority**
- **Token passing**


MAC
Media Access Control Address
00
1A
3F
F1
4C
C6
Organizationally Unique Identifier
Universally Administered Address

CSMA/CA

Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA)

Carrier sense multiple access with collision avoidance (CSMA/CA) is a media access control policy that regulates how data packets are transmitted between two computer nodes.

This method avoids collision by configuring each computer terminal to make a signal before transmission.

The signal is carried out by the transmitting computer to avoid a collision.

Multiple access implies that many computers are attempting to transmit data.

Collision avoidance means that when a computer node transmitting data states its intention, the other waits at a specific length of time before resending the data.

CSMA/CA is data traffic regulation is slow and adds cost in having each computer node signal its intention before transmitting data. It used only on Apple networks.

CSMA/CD

Carrier Sense Multiple Access with Collision Detection (CSMA/CD)

Carrier sense multiple access with collision detection (CSMA/CD) is the opposite of CSMA/CA.

Instead of detecting data to transmit signal intention to prevent a collision, it observes the cable to detect the signal before transmitting.

Collision detection means that when a collision is detected by the media access control policy, transmitting by the network stations stops at a random length of time before transmitting starts again.

It is faster than CSMA/CA as it functions in a network station that involves fewer data frames being transmitted.

CSMA/CD is not as efficient as CSMA/CA in preventing network collisions.

This is because it only detects huge data traffic in the network cable.

Huge data traffic increases the possibility of a collision taking place.

It is used on the Ethernet network.

Demand Priority

Demand Priority

The demand priority is an improved version of the Carrier sense multiple access with collision detection (CSMA/CD).

This data control policy uses an ‘active hub’ in regulating how a network is accessed.

Demand priority requires that the network terminals obtain authorization from the active hub before data can be transmitted.

Another distinct feature of this MAC control policy is that data can be transmitted between the two network terminals at the same time without collision.

In the Ethernet media, demand priority directs that data is transmitted directly to the receiving network terminal.

Token Passing

Token Passing

This media access control method uses free token passing to prevent a collision.

Only a computer that possesses a free token, which is a small data frame, is authorized to transmit.

Transmission occurs from a network terminal that has a higher priority than one with a low priority.

Token passing flourishes in an environment where a large number of short data frames are transmitted.

This media access control policy is highly efficient in avoiding a collision.

Token Passing

Token Passing

Possession of the free token is the only key to transmitting data by a network node.

Each terminal holds this free token for a specific amount of time if the network with the high priority does not have data to transmit, the token is passed to the adjoining station in the network.

Media access control regulates how a network is

accessed by computer terminals and transmits from one terminal to the other without collision.

This is achieved through CSMA/CD, CSMA/CA, demand priority, or Token passing.

6.4
Bridge, Hub, Switch

BRIDGE

Bridge is physical device typically a box of two port which connect two network at Data Link Layer.

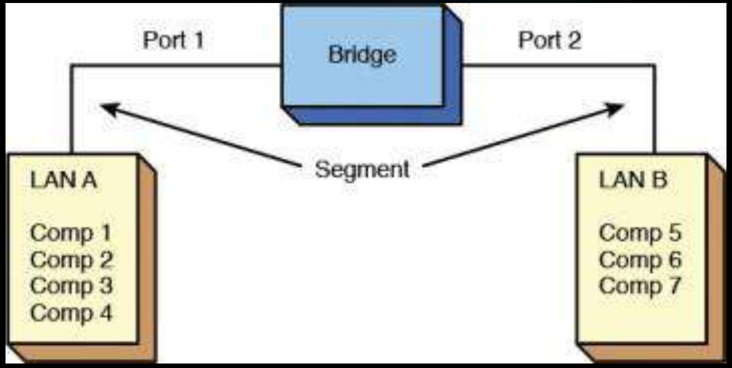
A Bridge provides packet filtering at Data link layer.

A bridge to join to existing LAN or two split one LAN in to two segment.

Bridge operate in promiscuous mode. Data packet enter the bridge through either one of the port and the bridge then read the destination address in each packet header and decides how to process that packet .

This is called packet filtering.

If the destination address of a packet arriving from one network segment is that of a computer on the other segment, the bridge Tx it out from other port.



If the destination address of a computer on a same network segment as the computer that generated it, the bridge discard the packet

BRIDGE & Collision

A collisions domain is a network that is constructed so that when two computers transmit packet at the same time a collision occurs.

When we add the new hub in existing network that the same collision domain as the original network because Hub relay the signal without filtering the packet

Bridge do not relay the signals to other network until they have receive the entire packet.

For this reason two computer on different side of the bridge do not cause to conflict.

Bridge maintain the internal address table that listed the hardware address of the computer on both segment .

BRIDGE & Collision

When bridge receive the packet and read the destination address DLL header.

It check the address against its lists.

If the address is associated with the segment other than that from which the packet arrived, the bridge relay it to that segment there are three type of bridge

- Local Bridge
- Translation Bridge
- Remote Bridge

BRIDGE & Collision

Local Bridge

Standard type of bridge use to connect network segment of same type and same location is called local bridge.

This is simplest type of bridge, it does not modify the data in packet.

It simply read the address in the data link layer protocol and pass the packet or discard it.

BRIDGE & Collision

Translation bridge

It is DLL device that connect network using different network media or different protocol.

This bridge is more complicated than local bridge.

The bridge can thus connect an Ethernet segment to FDDI (Fiber Distributed Data Interface) segment or connect two different type of Ethernet type segment such as (100BaseTx).

BRIDGE & Collision

Remote Bridge

Remote bridge is designed to connect two network segment at distance locations using some form of wide area network link.

The link can be a modem connection leased telephone line or any type of WAN technology.

The advantage of using a bridge in this manner is that you reduce the amount of traffic passing over the WAN link., which is usually far slower and more expensive than the local Network.

HUB

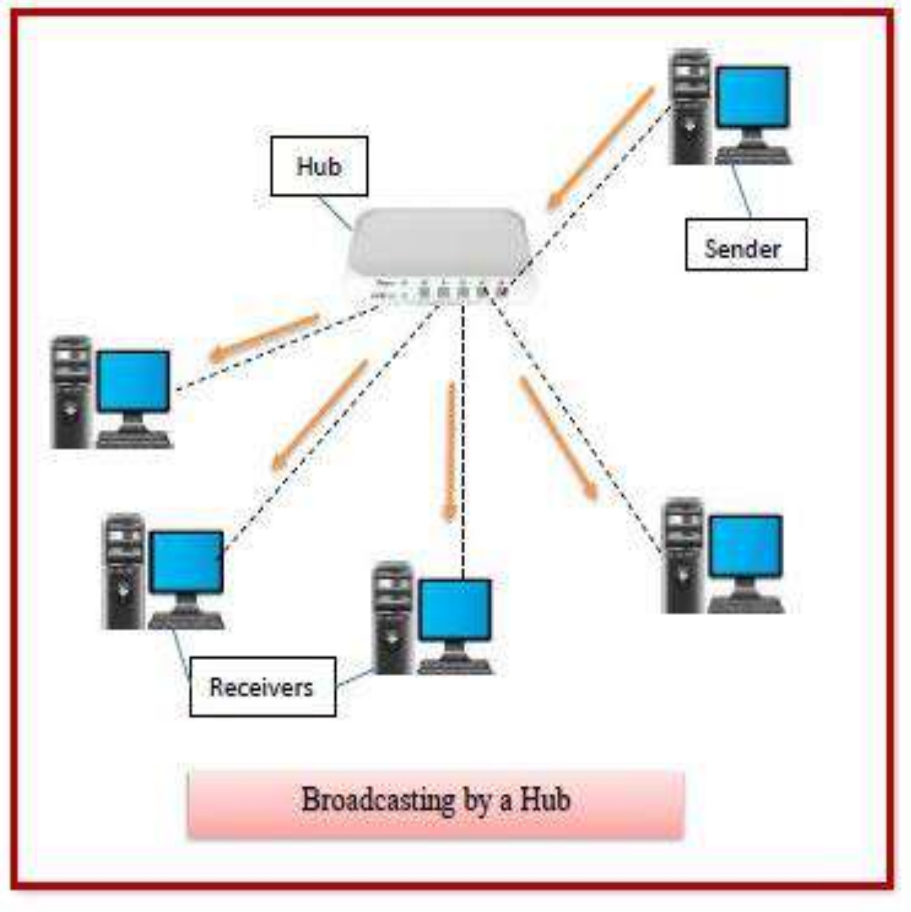
A hub is a common connection point, also known as a network hub, which is used for connection of devices in a network.

It works as a central connection for all the devices that are connected through a hub.

The hub has numerous ports.

If a packet reaches at one port, it is able to see by all the segments of the network due to a packet is copied to the other ports.

A network hub has no routing tables or intelligence (unlike a network switch or router), which is used to send information and broadcast all network data across each and every



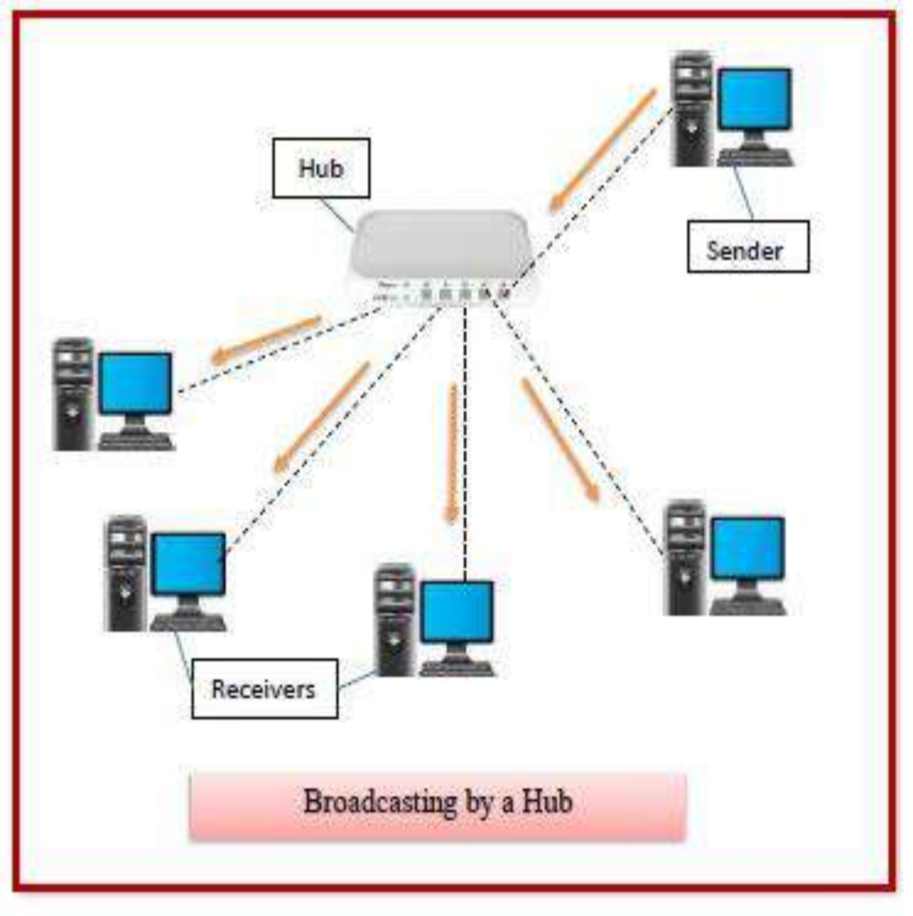
HUB

Although most of the hubs can recognize network troubles or errors like collisions, broadcasting all information to the several ports can be a security risk and cause bottlenecks.

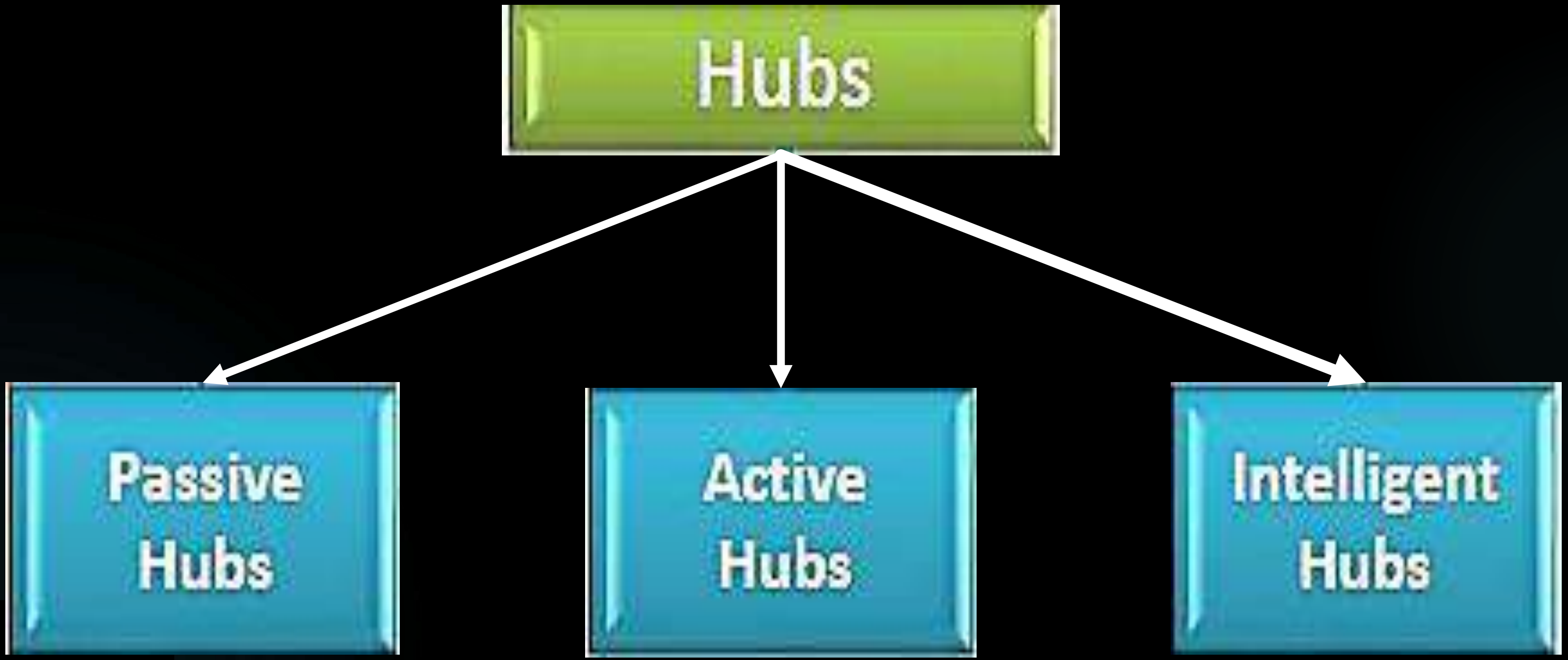
The network hubs were popular in the past time as they were cheaper as compared to a switch or router.

Nowadays, switches are much cheaper than a hub and provide a better solution for any network.

Furthermore, a hub is no IP address, as it is a dumb device



Types of HUB



Types of HUB

Passive HUB

The passive hubs are the connection point for wires that helps to make the physical network.

It is capable of determining the bugs and faulty hardware.

It accepts the packet over a port and circulates it to all ports.

It includes connectors (10base-2 port and RJ-45) that can be applied as a standard in your network.

This connector is connected to all local area network (LAN) devices.

Additionally, the advanced passive hubs have AUI ports, which are connected as the transceiver according to the network design

Types of HUB

Active HUB

As compared to a passive hub, it includes some additional features.

It is able to monitor the data sent to the connected devices.

It plays an important role between the connected devices with the help of store technology, where it checks the data to be sent and decides which packet to send first.

It has the ability to fix the damaged packets when packets are sending, and also able to hold the direction of the rest of the packets and distribute them.

If a port receives a weak signal, but still it is readable, then the active hub reconstructs the weak signal into a stronger signal before its sending to other ports.

It can boost the signal if any connecting device is not working in the network. Therefore, it helps to make the continuity of services in LAN

Types of HUB

Intelligent HUB

It is a little smarter than passive and active hubs.

These hubs have some kinds of management software that help to analyze the problem in the network and resolve them.

It is beneficial to expend the business in networking; the management can assign users that help to work more quickly and share a common pool efficiently by using intelligent hubs.

However, it offers better performance for the local area network.

Furthermore, with any physical device, if any problem is detected, it is able to detect this problem easily

Features of HUB

- It acts with shared bandwidth and broadcasting.
- It includes only one collision domain and broadcast domain.
- It works at the physical layer of the OSI model and also offers support for half-duplex transmission mode.
- Mainly packet collisions occur inside the hub.
- It also has a feature of flexibility, which means it includes a high transmission rate to different devices.
- A hub cannot filter data.
- It is a non-intelligent network device, as it sends message to all ports.
- It primarily broadcasts messages. So, the collision domain of all nodes connected through the hub stays one.
- Collisions may occurs during setup of transmission when more than one computers place data simultaneously in the corresponding ports.
- They generally have fewer ports of 4/12.

Application of HUB

The important applications of a hub are given below:

- Hub is used to create small home networks.
- It is used for network monitoring.
- They are also used in organizations to provide connectivity.
- It can be used to create a device that is available thought out of the network.

What HUB Do

Hubs work as a central connection between all network equipment and handle a data type, which is called frames.

If a frame is received, it is transmitted to the port of the destination computer after amplifying it.

A frame is passed to each of its ports in the hub, whether it is destined only for one port.

It does not include the way of deciding a frame to which port it should be sent.

Therefore, a frame has to transmit to every port, which ensures that it will reach its intended destination that generates a lot of traffic on the network and can be caused to damage the network.

The hub is slower as compared to standard switch as it is not able to send or receive information at the same time, but a switch is more costly than a hub

Advantages of HUB

- It provides support for different types of Network Media.
- It can be used by anyone as it is very cheap.
- It can easily connect many different media types.
- The use of a hub does not impact on the network performance.
- Additionally, it can expand the total distance of the network

Disadvantages of HUB

- It has no ability to choose the best path of the network.
- It does not include mechanisms such as collision detection.
- It does not operate in full-duplex mode and cannot be divided into the Segment.
- It cannot reduce the network traffic as it has no mechanism

SWITCH

A network switch is a device that operates at the Data Link layer of the OSI model—Layer 2.

It takes in packets being sent by devices that are connected to its physical ports and sends them out again, but only through the ports that lead to the devices the packets are intended to reach.

They can also operate at the network layer--Layer 3 where routing occurs.

Switches are a common component of networks based on ethernet, Fibre Channel, Asynchronous Transfer Mode (ATM), and InfiniBand, among others. In general, though, most switches today use Ethernet.

Types of SWITCH

Switches vary in size, depending on how many devices you need to connect in a specific area, as well as the type of network speed/bandwidth required for those devices.

In a small office or home office, a four- or eight-port switch usually suffices, but for larger deployments you generally see switches up to 128 ports.

The form factor of a smaller switch is an appliance that you can fit on a desktop, but switches are also rack-mountable for placement in a wiring closet or data center or server farm.

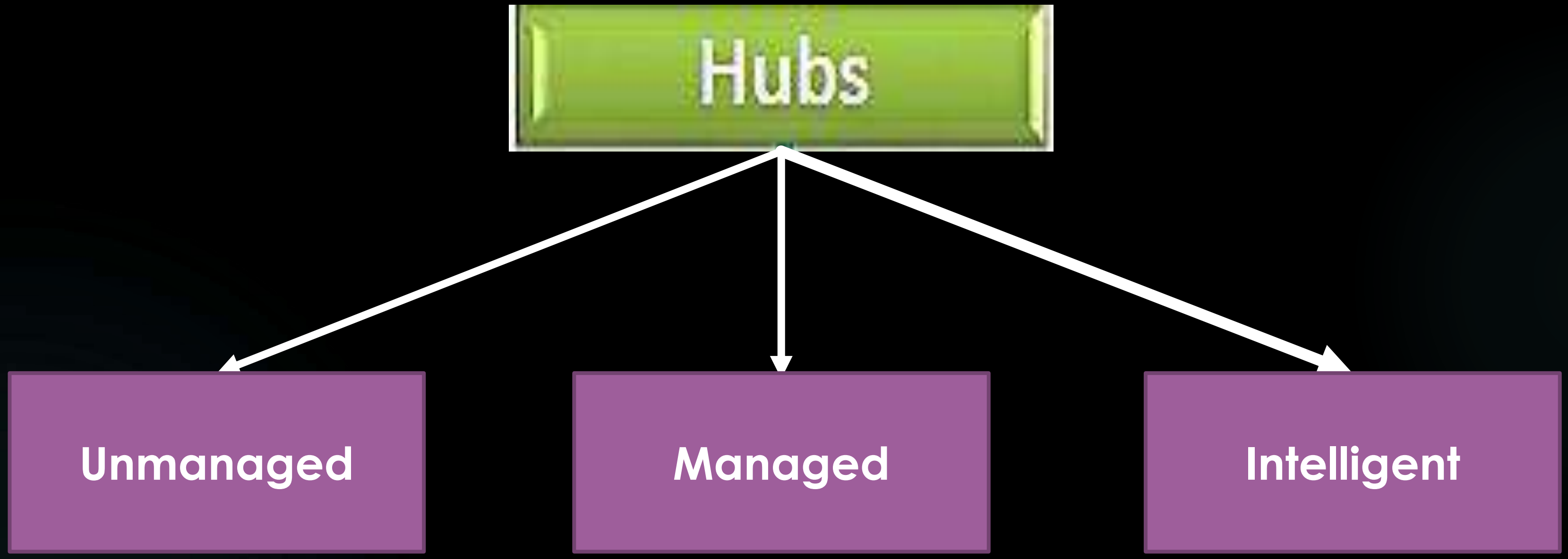
Sizes of rack-mountable switches range from 1U to 4U, but larger ones area also available.

Switches also vary in the network speed they offer, ranging from Fast ethernet (10/100 Mbps), Gigabit ethernet (10/100/1000 Mbps), 10 Gigabit (10/100/1000/10000 Mbps) and even 40/100 Gbps speeds.

Which speed to choose depends on the throughput needed for the tasks being supported

Types of SWITCH

Switches also differ in their capabilities. These are



Hubs

Unmanaged

Managed

Intelligent

Types of SWITCH

Unmanaged SWITCH

Unmanaged switches are the most basic, offering fixed configuration.

They are generally plug-and-play, which means they have few if any options for the user to choose from.

They may have default settings for features such as quality of service, but they cannot be changed.

The upside is that unmanaged switches are relatively inexpensive, but their lack of features make them unsuitable for most enterprise uses

Types of SWITCH

Managed SWITCH

Managed switches offer more functionality and features for IT professionals and are the type most likely seen in business or enterprise settings.

Managed switches have command-line interfaces (CLI) to configure them.

They support simple network management protocol (SNMP) agents that provide information that can be used to troubleshoot network problems.

They can also support virtual LANs, quality of service settings and IP routing.

The security is also better, protecting all types of traffic that they handle.

Because of their advanced features, managed switches cost much more than unmanaged switches

Types of SWITCH

Intelligent SWITCH

Smart or intelligent switches are managed switches that have some features beyond what an unmanaged switch offers, but fewer than a managed switch.

So they are more sophisticated than unmanaged switches, but they are also less expensive than a fully manageable switch.

They generally lack support for telnet access and have Web GUIs rather than CLIs.

Other options, such as VLANs, may not have as many features as those supported by fully managed switches.

But because they are less expensive, they may be a good fit for smaller networks with fewer financial resources and those with fewer feature needs

Management Feature of SWITCH

The full list of features and functionalities of a network switch will vary depending on the switch manufacturer and any additional software provided, but in general a switch will offer professionals the ability to:

- Enable and disable specific ports on the switch.
- Configure settings for duplex (half or full), as well as bandwidth.
- Set quality of service (QoS) levels for a specific port.
- Enable MAC filtering and other access control features.
- Set up SNMP monitoring of devices, including the health of the link.
- Configure port mirroring, for monitoring network traffic.

Difference between HUB and SWITCH

Hub	Switch
A hub works at the physical layer of the OSI model.	A switch works at the data link layer of the OSI model.
A hub contains a single domain of collision.	In switch, several ports include separate collision domains.
It performs frame flooding, which can be broadcast, unicast, or multicast.	It mainly performs broadcast, and also performs unicast and multicast when required.
In the hub, the transmission mode is Half-duplex	In switch, the transmission mode is full-duplex.
It uses electrical signal orbits.	It uses frame & packet.
It does not support the Spanning-Tree protocol.	It supports Multiple Spanning-Tree.

Difference between HUB and SWITCH

Hub	Switch
In the hub, mostly collisions occur in setup.	In full-duplex switch does not occur collisions.
It is a passive device.	It is an active device.
A hub is not capable of storing MAC addresses.	It uses accessible content memory, which can be accessed by application-specific integrated chips (ASIC).
It is not an intelligent device.	A switch is an intelligent device.
The speed of the hub network is up to 10 Mb per second.	The speed of switch is 10/100 Mbps, 1 Gbps, and 10 Gbps.

6.4
Ethernet (CSMA/CD), Fiber Channel

CARRIER SENSE MULTIPLE ACCESS

- It was developed to minimize the chance of collision and to increase the performance.
- It requires that each station first listen to the medium before sending.
- It is based on the principle sense before transmit or listen before talk.
- It can reduce the possibility of collision, but it cannot eliminate it.
- In CSMA a station senses the carrier on the channel before starting its own transmission.
- The vulnerable time for CSMA is the propagation time Tp.
- propagation time needed for a signal to propagate from one end of the medium to the other.
- When a station sends a frame, and any other station tries to send a frame during this time, a collision will result.
- When a channel is sense to be idle, a station can take one of the three different approaches to transmit a packet on to the channel. These three protocols are as follows:

CARRIER SENSE MULTIPLE ACCESS

Non-persistent CSMA

- In non-persistent CSMA, when a station having a packet to transmit and finds that the channel is busy, it backs off for a fixed interval of time.
- It then checks the channel again and if the channel is free then it transmits.
- The back-off delay is determined by the transmission of a frame, propagation time and other system parameter.
- If the channel is already in use, the station does not continuously sense it for the purpose of seizing it immediately upon detecting the end of the previous transmission.
- But it waits a random period of time and again check for activity

CARRIER SENSE MULTIPLE ACCESS

1-Persistent CSMA

- Any station wishing to transmit, monitor the channel continuously until the channel is idle and then transmit with probability one, hence the name 1-persistent.
- When two or more stations are waiting to transmit, a collision is guaranteed. Since each station will transmit immediately at the end of busy period.
- In this case each will wait a random amount of time and will then reattempt to transmit.
- As in the case with non-persistent CSMA, the performance of 1-persistent CSMA protocol depends only on the channel delay time.

CARRIER SENSE MULTIPLE ACCESS

P-Persistent CSMA:-

- To reduce the probability of collision in 1-persistent CSMA, not all allowed transmitting immediately, after the channel is idle.
- When a station becomes ready to send and its sense the channel to be idle, it either transmits with a probability p or it defers transmission by one time slot with a probability q=1-p.
- If the differed slot is also idle, the station either transmits with probability p or defers again with a probability q.
- This process is repeated until either packets are transmitted of the channel is busy

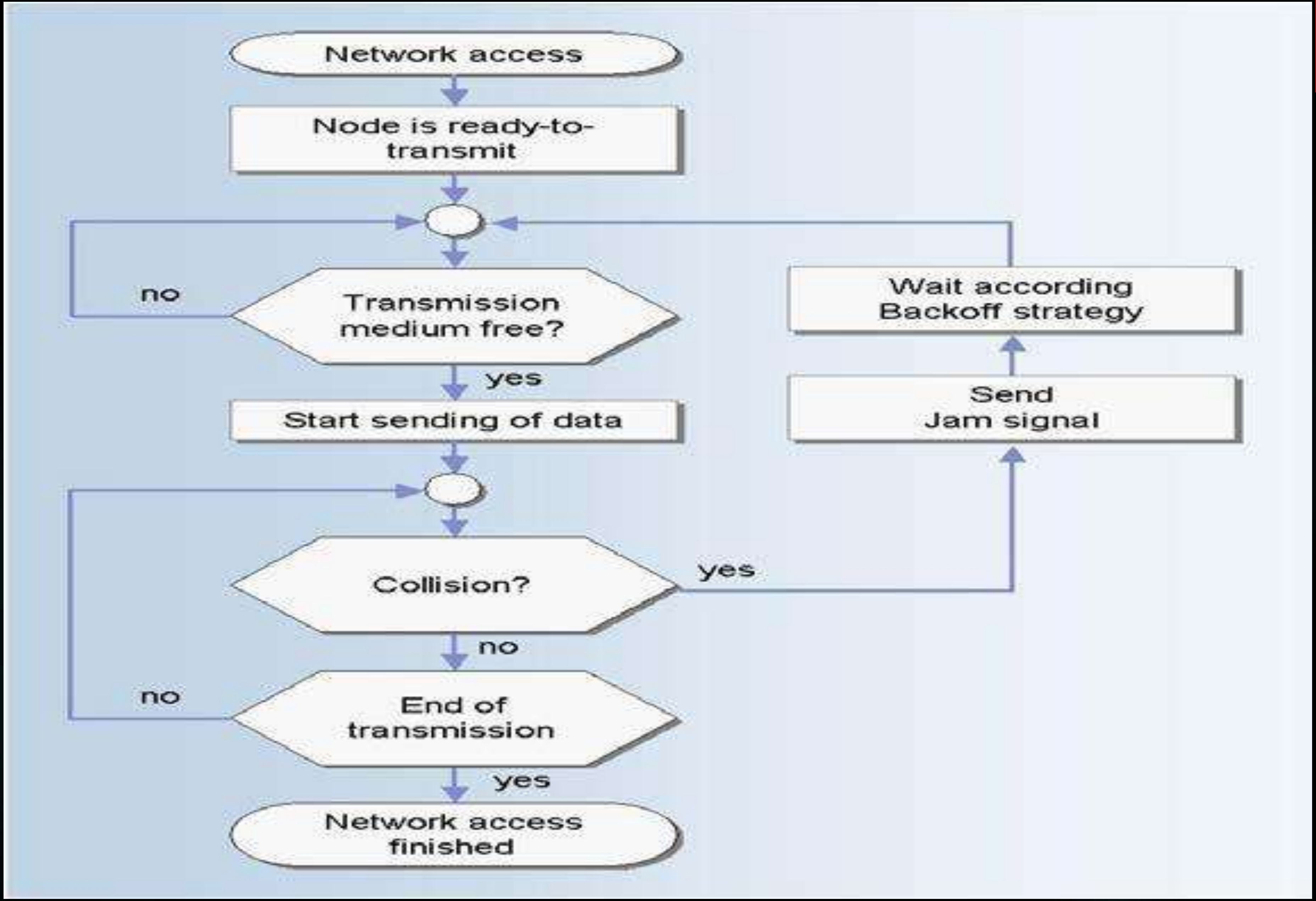
CSMA/CD

- CSMA/CD is the most commonly used protocol for LANs.
- CSMA/CD specifications were developed jointly by digital equipment corporation (DEC), Intel, and Xerox. This network is called as Ethernet. The IEEE802.3
- CSMA/CD stands for LAN are based on Ethernet specification.
- The basic protocol is that, a station with a message to send must monitor the channel to see if any other station is sending.
- If another station is sending, the second station must wait or defer, until the sending station has finished.
- Then it may send its message. If no station was sending at the time that it first listened, the station may send its message immediately.
- The term carrier sense indicates this listening before transmitting behavior.
- **Carrier Sense Multiple Access/Collision Detect** (CSMA/CD) is the protocol for carrier transmission access in Ethernet networks

CSMA/CD

- **Carrier-sense multiple access with collision detection** describes how the Ethernet protocol regulates communication among nodes.
- On Ethernet, any station can send a frame at any time.
- Each station senses whether the medium is idle and therefore available for use.
- If it is, the station begins to transmit its first frame.
- If another station also tries to transmit at the same time, a collision occurs and the frames are discarded and then a jamming signal is sent throughout the network in order to notify all stations of the collision.
- Each station then waits for a random period of time and retries.
- If another collision occurs, the time intervals from which the random waiting time is selected are increased step by step.
- This is known as exponential back off. The stations retry until successful transmission of the frame.CSMA/CD is specified in the IEEE 802.3 standard.

CSMA/CD


Network access
Node is ready-to-transmit
no
Transmission medium free?
yes
Start sending of data
Wait according Backoff strategy
Send Jam signal
Collision?
yes
no
no
End of transmission
yes
Network access finished

CSMA/CD

- ❑ The **jam signal** is a signal that carries a 32-bit binary pattern sent by a data station to inform the other stations that they must not transmit.
- ❑ **Transmit** – The stations (nodes) sends the frames to other stations (nodes)
- ❑ **Carrier Sense** – The stations (nodes) listen to the medium if it is idle for transmission
- ❑ **Back off** – After collision occurs, a jam signal is sent to notify all stations of the collision.
- ❑ After the jam signal is sent, the stations (nodes) wait for a random period of time called Back off period.
- ❑ A protocol is required for transmitting station to monitor the channel while sending each of its messages and to detect such collisions.

CSMA/CD

- ❑ If two or more stations have message to send at the same time and they are separated by significant distances on the bus/channel.
- ❑ Each may begin transmitting at roughly the same time without being aware of the other station.
- ❑ The signal from each station will superimpose on the channel and is garbled beyond the decoding ability of the receiving station.
- ❑ This is termed as collision .
- ❑ When a collision has been detected, each of sending stations must cease transmitting, wait for a random length of time, and then try again.
- ❑ Because of quick termination of transmission time and bandwidth is saved. Therefore CSMA/CD is more efficient than ALOHA and CSMA.
- ❑ CSMA/CD network work best on a bus, multipoint topology with busty asynchronous transmission.
- ❑ CSMA/CD has totally decentralized control and is based on connection access.

CSMA/CA

CSMA with collision avoidance (CSMA/CA)

Collision avoidance is used to improve the performance of the CSMA method by attempting to divide the channel somewhat equally among all transmitting nodes within the collision domain.

1. Carrier Sense

- Prior to transmitting, a node first listens to the shared medium (such as listening for wireless signals in a wireless network) to determine whether another node is transmitting or not.

Collision Avoidance

- If another node was heard, we wait for a period of time for the node to stop transmitting before listening again for a free communications channel.

CSMA/CA

- Request to Send/Clear to Send (RTS/CTS) may optionally be used at this point to mediate access to the shared medium.
- This goes some way to alleviating the problem of hidden nodes because, for instance, in a wireless network, the Access Point only issues a Clear to send to one node at a time.
- **Transmission**: if the medium was identified as being clear or the node received CTS to explicitly indicate it can send, it sends the frame in its entirety.
- Unlike CSMA/CD, it is very challenging for a wireless node to listen at the same time as it transmits (its transmission will dwarf any attempt to listen).

CSMA/CA

- Although CSMA/CA has been used in a variety of wired communication systems, it is particularly beneficial in a wireless LAN due to a common problem of multiple stations being able to see the Access Point, but not each other.
- This is due to differences in transmitting power, and receives sensitivity.
- CSMA/CA performance is based largely upon the modulation technique used to transmit the data between nodes.
- Studies show that under ideal propagation conditions (simulations), Direct Sequence Spread Spectrum (DSS) provides the highest throughput for all nodes on a network when used in conjunction with CSMA/CA and the IEEE 802.11 RTS/CTS exchange under light network load conditions

6.5
Wireless LAN Technology

WIRELESS LAN

Wireless LAN stands for **Wireless Local Area Network**.

It is also called LAWN (**Local Area Wireless Network**).

Wireless LANs (WLANs) are wireless computer networks that use high-frequency radio waves instead of cables for connecting the devices within a limited area forming LAN (Local Area Network).

Users connected by wireless LANs can move around within this limited area such as home, school, campus, office building, railway platform, etc.

The IEEE 802.11 group of standards defines the technologies for wireless LANs.

For path sharing, 802.11 standard uses the Ethernet protocol and CSMA/CA (carrier sense multiple access with collision avoidance).

It also uses an encryption method i.e. wired equivalent privacy algorithm.

WIRELESS LAN

Wireless LANs provide high speed data communication in small areas such as building or an office.

WLANs allow users to move around in a confined area while they are still connected to the network.

In some instance wireless LAN technology is used to save costs and avoid laying cable, while in other cases, it is the only option for providing high-speed internet access to the public.

Whatever the reason, wireless solutions are popping up everywhere.

TYPES WIRELESS LAN

Types of WLANS

WLANs, as standardized by IEEE 802.11, operates in two basic modes, infrastructure, and ad hoc mode.

Infrastructure Mode – Mobile devices or clients connect to an access point (AP) that in turn connects via a bridge to the LAN or Internet.

The client transmits frames to other clients via the AP.

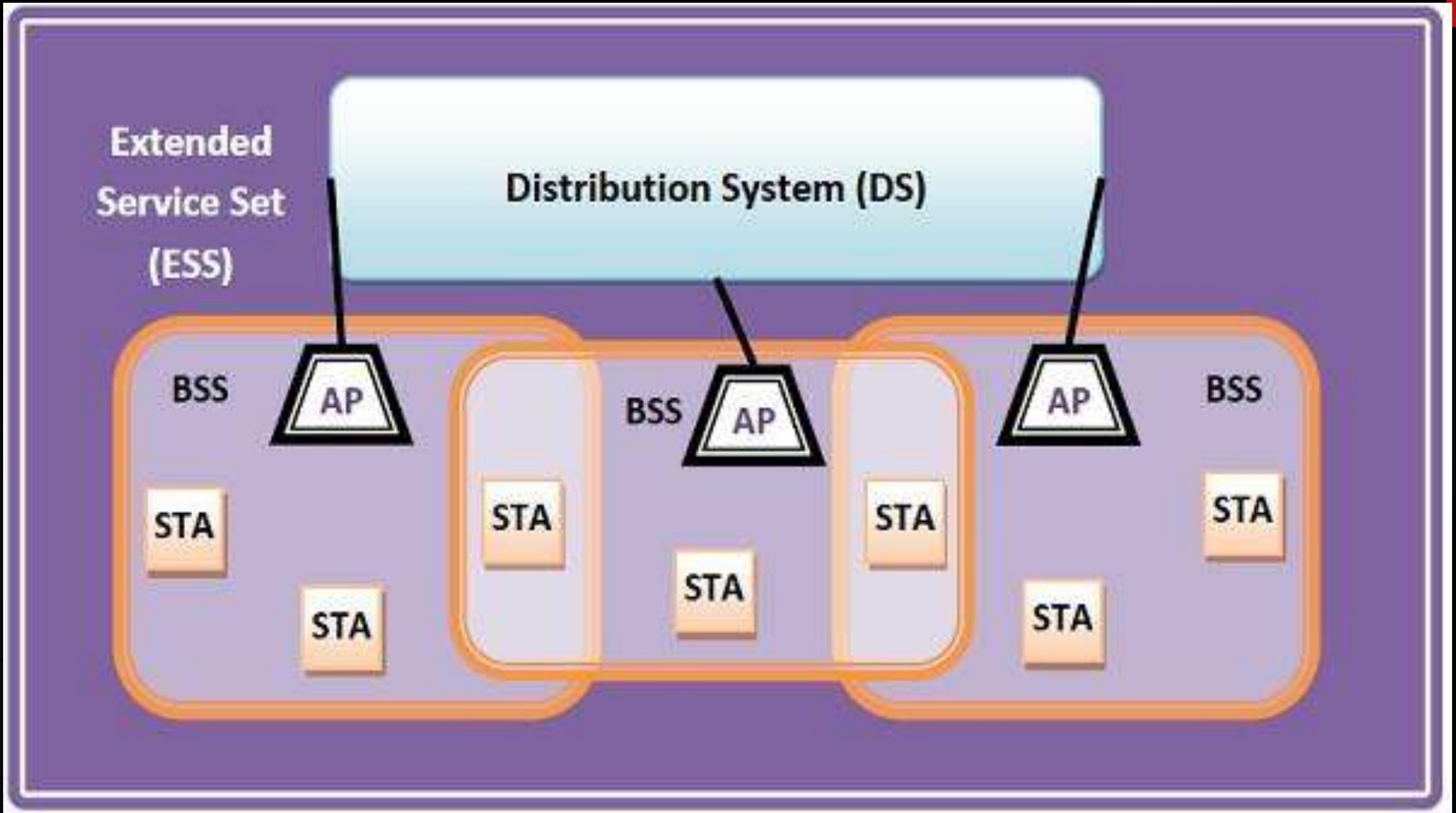
Ad Hoc Mode – Clients transmit frames directly to each other in a peer-to-peer fashion

Components of WLANs

The components of WLAN architecture as laid down in IEEE 802.11 are –

- **Stations (STA)** – Stations comprises of all devices and equipment that are connected to the wireless LAN. Each station has a wireless network interface controller. A station can be of two types –
  - Wireless Access Point (WAP or AP)
  - Client
- **Basic Service Set (BSS)** – A basic service set is a group of stations communicating at the physical layer level. BSS can be of two categories –
  - Infrastructure BSS
  - Independent BSS
- **Extended Service Set (ESS)** – It is a set of all connected BSS.
- **Distribution System (DS)** – It connects access points in ESS.

WIRELESS LAN



Advantages of Wireless LAN

- They provide clutter-free homes, offices and other networked places.
- The LANs are scalable in nature, i.e. devices may be added or removed from the network at greater ease than wired LANs.
- The system is portable within the network coverage. Access to the network is not bounded by the length of the cables.
- **Flexibility:** Within radio coverage, nodes can communicate without further restriction. Radio waves can penetrate walls, senders and receivers can be placed anywhere (also non-visible, e.g., within devices, in walls etc.).
- **Planning:** Only wireless ad-hoc networks allow for communication without previous planning, any wired network needs wiring plans.

Advantages of Wireless LAN

- **Design:** Wireless networks allow for the design of independent, small devices which can for example be put into a pocket. Cables not only restrict users but also designers of small notepads, PDAs, etc.
- **Robustness:** Wireless networks can handle disasters, e.g., earthquakes, flood etc. whereas, networks requiring a wired infrastructure will usually break down completely in disasters.
- **Cost:** The cost of installing and maintaining a wireless LAN is on average lower than the cost of installing and maintaining a traditional wired LAN, for two reasons. First, after providing wireless access to the wireless network via an access point for the first user, adding additional users to a network will not increase the cost. And second, wireless LAN eliminates the direct costs of cabling and the labor associated with installing and repairing it.
- **Ease of Use:** Wireless LAN is easy to use and the users need very little new information to take advantage of WLANs.

Disadvantages of Wireless LAN

- Since radio waves are used for communications, the signals are noisier with more interference from nearby systems.
- Greater care is needed for encrypting information. Also, they are more prone to errors. So, they require greater bandwidth than the wired LANs.
- WLANs are slower than wired LANs.
- **Quality of Services:** Quality of wireless LAN is typically lower than wired networks. The main reason for this is the lower bandwidth due to limitations is radio transmission, higher error rates due to interference and higher delay/delay variation due to extensive error correction and detection mechanisms.
- **Proprietary Solutions:** Due to slow standardization procedures, many companies have come up with proprietary solutions offering standardization functionality plus many enhanced features. Most components today adhere to the basic standards IEEE 802.11a or 802.11b.

Disadvantages of Wireless LAN

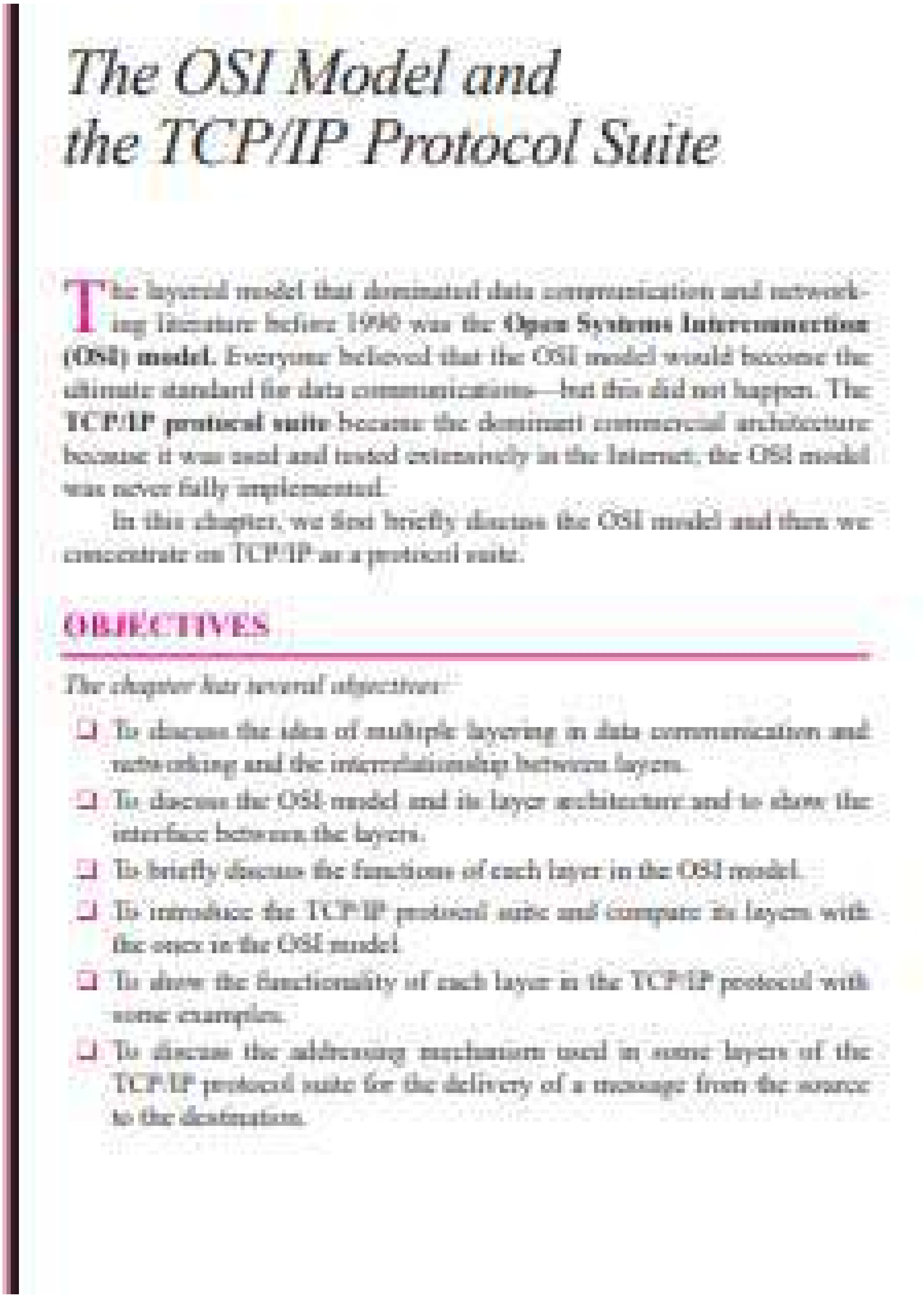
- **Restrictions:** Several govt. and non-govt. institutions world-wide regulate the operation and restrict frequencies to minimize interference.
- **Global operation:** Wireless LAN products are sold in all countries so, national and international frequency regulations have to be considered.
- **Low Power:** Devices communicating via a wireless LAN are typically power consuming, also wireless devices running on battery power. Whereas the LAN design should take this into account and implement special power saving modes and power management functions.
- **License free operation:** LAN operators don't want to apply for a special license to be able to use the product. The equipment must operate in a license free band, such as the 2.4 GHz ISM band.
- **Robust transmission technology:** If wireless LAN uses radio transmission, many other electrical devices can interfere with them (such as vacuum cleaner, train engines, hair dryers, etc.).Wireless LAN transceivers cannot be adjusted for perfect transmission is a standard office or production environment.

Thank You

DCCN (UNIT-7)

OSI Model and TCP/IP protocol suite

By : Behrouz A. Forouzan



The OSI Model and the TCP/IP Protocol Suite

The layered model that dominated data communication and networking literature before 1990 was the **Open Systems Interconnection (OSI) model.** Everyone believed that the OSI model would become the ultimate standard for data communications—but this did not happen. The **TCP/IP protocol suite** became the dominant commercial architecture because it was used and tested extensively in the Internet; the OSI model was never fully implemented.

In this chapter, we first briefly discuss the OSI model and then we concentrate on TCP/IP as a protocol suite.

OBJECTIVES

The chapter has several objectives:

❑ To discuss the idea of multiple layering in data communication and networking and the interrelationship between layers.

❑ To discuss the OSI model and its layer architecture and to show the interface between the layers.

❑ To briefly discuss the functions of each layer in the OSI model.

❑ To introduce the TCP/IP protocol suite and compare its layers with the ones in the OSI model.

❑ To show the functionality of each layer in the TCP/IP protocol with some examples.

❑ To discuss the addressing mechanism used in some layers of the TCP/IP protocol suite for the delivery of a message from the source to the destination.

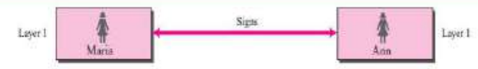
PROTOCOL LAYERS

In previous Chapter, we discussed that a protocol is required when two entities need to communicate. When communication is not simple, we may divide the complex task of communication into several layers. In this case, we may need several protocols, one for each layer.

Let us use a scenario in communication in which the role of protocol layering may be better understood. We use two examples. In the first example, communication is so simple that it can occur in only one layer. In the second example, we need three layers.

Example 2.1 Assume Maria and Ann are neighbors with a lot of common ideas. However, Maria speaks only Spanish, and Ann speaks only English. Since both have learned the sign language in their child hood, they enjoy meeting in a cafe a couple of days per week and exchange their ideas using signs. Occasionally, they also use a bilingual dictionary. Communication is face to face and hap pens in one layer as shown in Figure 2.1.

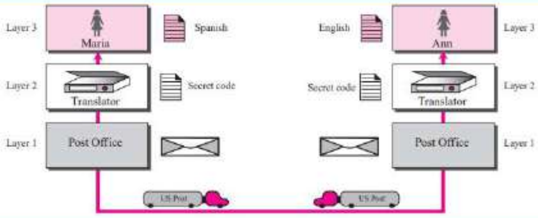
Figure 2.1 *Example 2.1*



Example 2.2 Now assume that Ann has to move to another town because of her job. Before she moves, the two meet for the last time in the same cafe. Although both are sad, Maria surprises Ann when she opens a packet that contains two small machines. The first machine can scan and transform a letter in English to a secret code or vice versa. The other machine can scan and translate a letter in Spanish to the same secret code or vice versa. Ann takes the first machine; Maria keeps the second one. The two friends can still communicate using the secret code, as shown in Figure 2.2.

Communication between Maria and Ann happens as follows. At the third layer, Maria writes a letter in Spanish, the language she is comfortable with. She then uses the translator machine that scans the letter and creates a letter in the secret code. Maria then puts the letter in an envelop and drops it to the post office box. The letter is carried by the post office truck to the post office of the city where Ann lives now. In the post office, the letter is delivered to the Ann residence. Ann uses her own machine to change the secret code to a letter in the English language. The commu nication from Ann to Maria uses the same process, but in the reverse direction. The communica tion in both directions is carried in the secret code, a language that neither Maria nor Ann understands, but through the layered communication, they can exchange ideas.

Figure 2.2 *Example 2.2*



Hierarchy

Using Example 2.2, there are three different activities at the sender site and another three activities at the receiver site. The task of transporting the letter between the sender and the receiver is done by the carrier. Something that is not obvious immediately is that the tasks must be done in the order given in the hierarchy. At the sender site, the letter must be written, translated to secret code, and dropped in the mailbox before being picked up by the letter carrier and delivered to the post office. At the receiver site, the letter must be dropped in the recipient mailbox before being picked up and read by the recipient.

Services

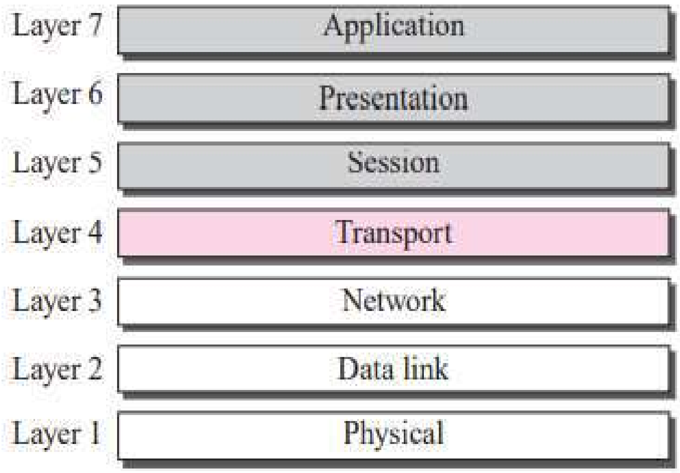
Each layer at the sending site uses the services of the layer immediately below it. The sender at the higher layer uses the services of the middle layer. The middle layer uses the services of the lower layer. The lower layer uses the services of the carrier.

THE OSI MODEL

Established in 1947, the International Standards Organization (ISO) is a multinational body dedicated to worldwide agreement on international standards. Almost three-fourths of countries in the world are represented in the ISO. An ISO standard that covers all aspects of network communications is the Open Systems Interconnection (OSI) model. It was first introduced in the late 1970s.

An open system is a set of protocols that allows any two different systems to communicate regardless of their underlying architecture. The purpose of the OSI model is to show how to facilitate communication between different systems without requiring changes to the logic of the underlying hardware and software. 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 was intended to be the basis for the creation of the protocols in the OSI stack.

The OSI model is a layered framework for the design of network systems that allows communication between all types of computer systems. It consists of seven separate but related layers, each of which defines a part of the process of moving information across a network (see Figure 2.3). Understanding the fundamentals of the OSI model provides a solid basis for exploring data communications.

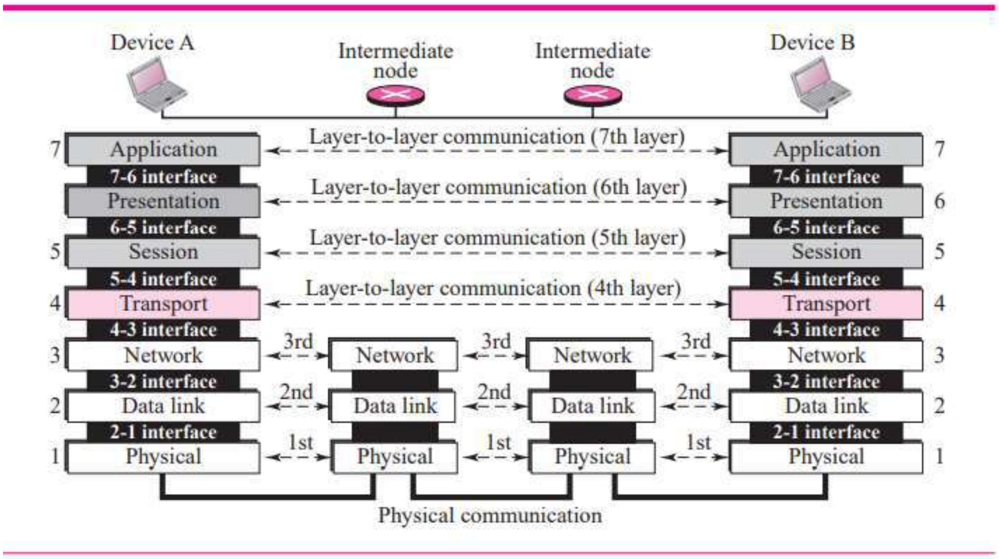


Layered Architecture

The OSI model is composed of seven ordered layers: physical (layer 1), data link (layer 2), network (layer 3), transport (layer 4), session (layer 5), presentation (layer 6), and application (layer 7). Figure 2.4 shows the layers involved when a message is sent from device A to device B. As the message travels from A to B, it may pass through many intermediate nodes. These intermediate nodes usually involve only the first three layers of the OSI model.

In developing the model, the designers distilled the process of transmitting data to its most fundamental elements. They identified which networking functions had related uses and collected those functions into discrete groups that became the layers. Each layer defines a family of functions distinct from those of the other layers. By defining and localizing functionality in this fashion, the designers created an architecture that is both comprehensive and flexible. Most important, the OSI model allows complete interoperability between otherwise incompatible systems.

Within a single machine, each layer calls upon the services of the layer just below it. Layer 3, for example, uses the services provided by layer 2 and provides services for layer 4. Between machines, layer x on one machine logically communicates with layer x on another machine. This communication is governed by an agreed-upon series of rules and conventions called protocols.



Layer-to-Layer Communication

In Figure 2.4, device A sends a message to device B (through intermediate nodes). At the sending site, the message is moved down from layer 7 to layer 1. At layer 1 the entire package is converted to a form that can be transferred to the receiving site. At the receiving site, the message is moved up from layer 1 to layer 7.

Interfaces between Layers

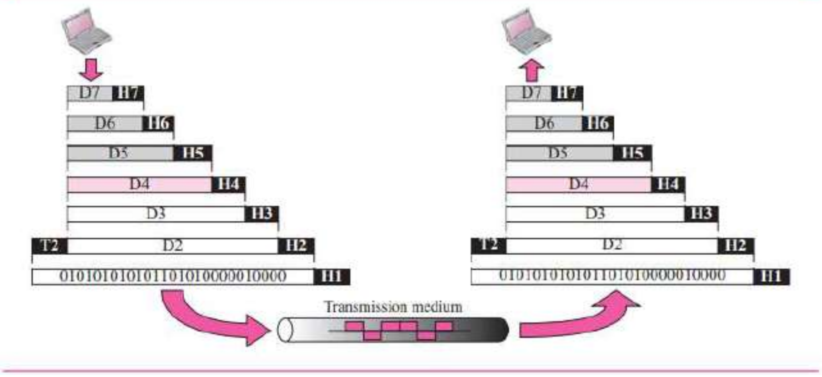
The passing of the data and network information down through the layers of the sending device and back up through the layers of the receiving device is made possible by an interface between each pair of adjacent layers. Each interface defines what information and services a layer must provide for the layer above it. Well-defined interfaces and layer functions provide modularity to a network. As long as a layer provides the expected services to the layer above it, the specific implementation of its functions can be modified or replaced without requiring changes to the surrounding layers.

Organization of the Layers

The seven layers can be thought of as belonging to three subgroups. Layers 1, 2, and 3—physical, data link, and network—are the network support layers; they deal with the physical aspects of moving data from one device to another (such as electrical specifications, physical connections, physical addressing, and transport timing and reliability). Layers 5, 6, and 7—session, presentation, and application—can be thought of as the user support layers; they allow interoperability among unrelated software systems. Layer 4, the transport layer, links the two subgroups and ensures that what the lower layers have transmitted is in a form that the upper layers can use. The upper OSI layers are almost always implemented in software; lower layers are a combination of hardware and software, except for the physical layer, which is mostly hardware.

In Figure 2.5, which gives an overall view of the OSI layers, D7 data means the data unit at layer 7, D6 data means the data unit at layer 6, and so on. The process starts at layer 7 (the application layer), then moves from layer to layer in descending, sequential order. At each layer, a header can be added to the data unit. At layer 2, a trailer may also be added. When the formatted data unit passes through the physical layer (layer 1), it is changed into an electromagnetic signal and transported along a physical link.

Figure 2.5 *An exchange using the OSI model*



Upon reaching its destination, the signal passes into layer 1 and is transformed back into digital form. The data units then move back up through the OSI layers. As each block of data reaches the next higher layer, the headers and trailers attached to it at the corresponding sending layer are removed, and actions appropriate to that layer are taken. By the time it reaches layer 7, the message is again in a form appropriate to the application and is made available to the recipient.

Encapsulation

Figure 2.5 reveals another aspect of data communications in the OSI model: encapsulation. A packet at level 7 is encapsulated in the 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 part of a packet at level N is carrying the whole packet (data and overhead) from level N + 1. The concept is called encapsulation because level N is not aware what part of the encapsulated packet is data and what part is the header or trailer. For level N, the whole packet coming from level N + 1 is treated as one integral unit.

Layers in the OSI Model

In this section we briefly describe the functions of each layer in the OSI model.

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 media. It also defines the procedures and functions that physical devices and interfaces have to perform for transmission to occur.

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

The physical layer is also concerned with the following:

❑ **Physical characteristics of interfaces and media.** The physical layer defines the characteristics of the interface between the devices and the transmission media. It also defines the type of transmission media (see Chapter 3).

❑ **Representation of bits.** The physical layer data consists of a stream of bits (sequence of 0s or 1s) with no interpretation. To be transmitted, bits must be encoded into signals—electrical or optical. The physical layer defines the type of encoding (how 0s and 1s are changed to signals).

❑ **Data rate.** The transmission rate—the number of bits sent each second—is also defined by the physical layer. In other words, the physical layer defines the duration of a bit, which is how long it lasts.

❑ **Synchronization of bits.** The sender and receiver must not only use the same bit rate but must also be synchronized at the bit level. In other words, the sender and the receiver clocks must be synchronized.

❑ **Line configuration.** The physical layer is concerned with the connection of devices to the media. In a point-to-point configuration, two devices are connected together through a dedicated link. In a multipoint configuration, a link is shared between several devices.

❑ **Physical topology.** The physical topology defines how devices are connected to make a network. Devices can be connected using a mesh topology (every device connected to every other device), a star topology (devices are connected through a central device), a ring topology (each device is connected to the next, forming a ring), or a bus topology (every device on a common link).

❑ **Transmission mode.** The physical layer also defines the direction of transmission between two devices: simplex, half-duplex, or full-duplex. In the simplex mode, only one device can send; the other can only receive. The simplex mode is a one way communication. In the half-duplex mode, two devices can send and receive, but not at the same time. In a full-duplex (or simply duplex) mode, two devices can send and receive at the same time.

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). 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 connecting device that connects the network to the next one.

❑ **Flow control**. If the rate at which the data is absorbed by the receiver is less than the rate produced at the sender, the data link layer imposes a flow control mechanism to prevent 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

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 (link), 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. 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 together 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.

Transport Layer

The transport layer is responsible for process-to-process delivery of the entire message. A process is an application program running on the 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. 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 four layers (physical, data link, network and transport) are not sufficient for some processes. The session layer is the network dialog controller. It establishes, maintains, and synchronizes the interaction between communicating systems. 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 (synchronization points) into a stream of data. For example, if a system is sending a file of 2,000 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.

Presentation Layer

The presentation layer is concerned with the syntax and semantics of the information exchanged between two systems. 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 should 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 assure 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. 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 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 you to log on.

❑ **File transfer, access, and management (FTAM).** 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.

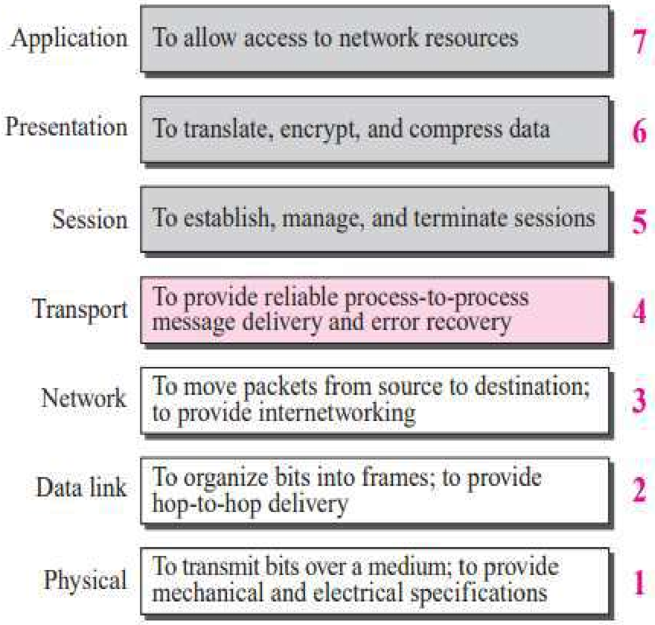
❑ **E-mail services.** This application provides the basis for e-mail forwarding and storage.

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

Summary of OSI Layers

Figure 2.6 shows a summary of duties for each layer. In the next section, we describe how some of these duties are mixed and spread into five categories in the TCP/IP proto col suite.

Figure 2.6 *Summary of OSI layers*



Application	To allow access to network resources	7
Presentation	To translate, encrypt, and compress data	6
Session	To establish, manage, and terminate sessions	5
Transport	To provide reliable process-to-process message delivery and error recovery	4
Network	To move packets from source to destination; to provide internetworking	3
Data link	To organize bits into frames; to provide hop-to-hop delivery	2
Physical	To transmit bits over a medium; to provide mechanical and electrical specifications	1

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 match exactly with those in the OSI model. The original TCP/IP protocol suite was defined as four software layers built upon the hardware. Today, however, TCP/IP is thought of as a five-layer model with the layers named similarly to the ones in the OSI model. Figure 2.7 shows both configurations.

Comparison between OSI and TCP/IP Protocol Suite

When we compare the two models, we find that two layers, session and presentation, are missing from the TCP/IP protocol suite. These two layers were not added to the TCP/IP protocol suite after the publication of the OSI model. The application layer in the suite is usually considered to be the combination of three layers in the OSI model, as shown in Figure 2.8.

Figure 2.7 *Layers in the TCP/IP Protocol Suite*

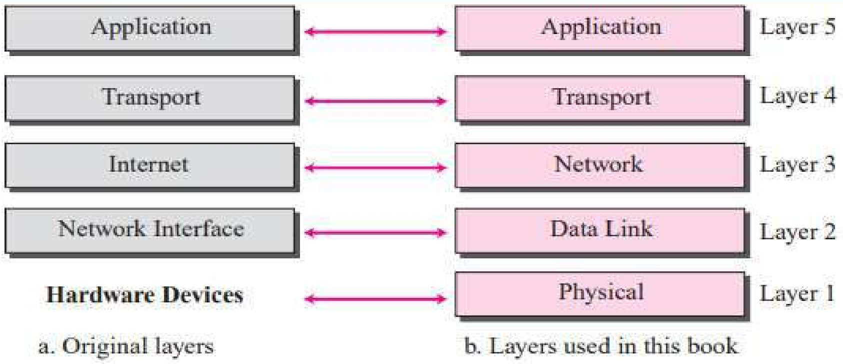
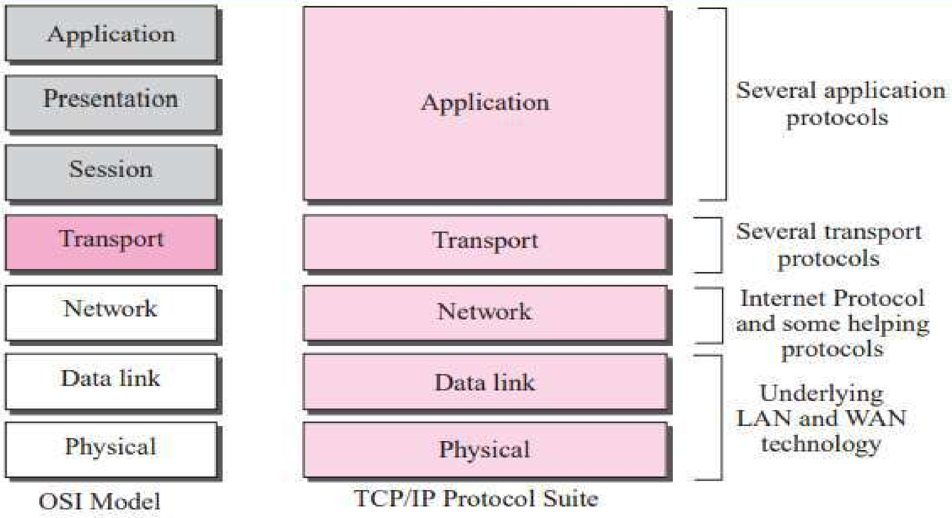


Figure 2.8 *TCP/IP and OSI model*



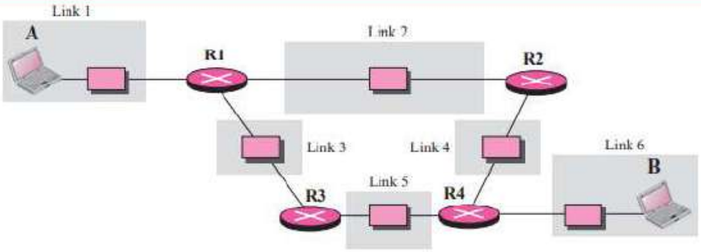
Two reasons were mentioned for this decision. First, TCP/IP has more than one transport-layer protocol. Some of the functionalities of the session layer are available in some of the transport layer protocols. Second, the application layer is not only one piece of software. Many applications can be developed at this layer. If some of the functionalities mentioned in the session and presentation are needed for a particular application, it can be included in the development of that piece of software.

TCP/IP is a hierarchical protocol made up of interactive modules, each of which provides a specific functionality, but 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.

Layers in the TCP/IP Protocol Suite

In this section, we briefly discuss the purpose of each layer in the TCP/IP protocol suite. When we study the purpose of each layer, it is easier to think of a private internet, instead of the global Internet. We assume that we want to use the TCP/IP suite in a small, private internet. Such an internet is made up of several small networks, which we call links. A link is a network that allows a set of computers to communicate with each other. For example, if all computers in an office are wired together, the connection makes a link. If several computers belonging to a private company are connected via a satellite channel, the connection is a link. A link, as we discussed in Chapter 3, can be a LAN (local area network) serving a small area or a WAN (wide area network) serving a larger area. We also assume that different links are connected together by devices called rout ers or switches that route the data to reach their final destinations. Figure 2.9 shows our imaginary internet that is used to show the purpose of each layer. We have six links and four routers (R1 to R4). We have shown only two computers, A and B.

Figure 2.9 *A private internet*

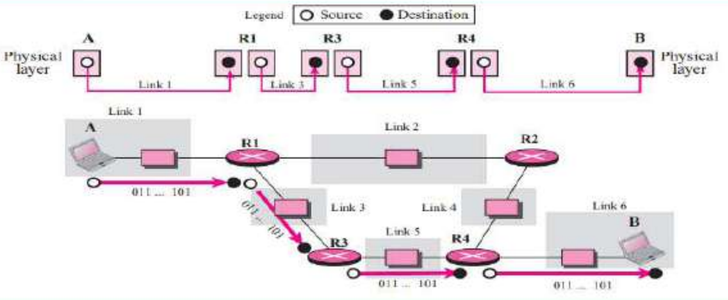


Physical Layer

TCP/IP does not define any specific protocol for the physical layer. It supports all of the standard and proprietary protocols. At this level, the communication is between two hops or nodes, either a computer or router. The unit of communication is a single bit. When the connection is established between the two nodes, a stream of bits is flowing between them. The physical layer, however, treats each bit individually. Figure 2.10 shows the communication between nodes. We are assuming that at this moment the two computers have discovered that the most efficient way to communicate with each other is via routers R1, R3, and R4. How this decision is made is the subject of some future chapters.

Note that if a node is connected to n links, it needs n physical-layer protocols, one for each link. The reason is that different links may use different physical-layer protocols. The figure, however, shows only physical layers involved in the communication. Each computer involves with only one link; each router involves with only two links. As Figure 2.10 shows, the journey of bits between computer A and computer B is made of four independent short trips. Computer A sends each bit to router R1 in the format of the protocol used by link 1. Router 1 sends each bit to router R3 in the format dictated by the protocol used by link 3. And so on. Router R1 has two three physical layers (two are shown in our scenario). The layer connected to link 1 receives bits according to the format of the protocol used by link 1; the layer connected to link 3 sends bits according to the format of the protocol used by link 3. It is the same situation with the other two routers involved in the communication.

Figure 2.10 *Communication at the physical layer*



The unit of communication at the physical layer is a bit.

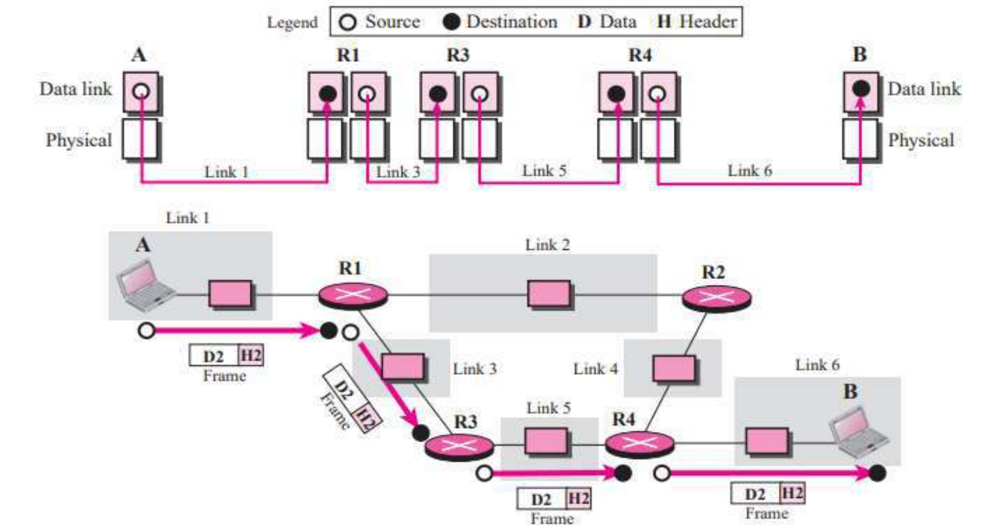
The responsibility of the physical layer, in addition to delivery of bits, matches with what mentioned for the physical layer of the OSI model, but it mostly depends on the underlying technologies that provide links. We see in the next chapter that they are, for example, many protocols for the physical layer of LANs or WANs.

Data Link Layer

TCP/IP does not define any specific protocol for the data link layer either. It supports all of the standard and proprietary protocols. At this level, the communication is also between two hops or nodes. The unit of communication however, is a packet called a frame. A frame is a packet that encapsulates the data received from the network layer with an added header and sometimes a trailer. The head, among other communication information, includes the source and destination of frame. The destination address is needed to define the right recipient of the frame because many nodes may have been connected to the link. The source address is needed for possible response or acknowledgment as may be required by some protocols.

Figure 2.11 shows the communication at the data link layer.

Figure 2.11 *Communication at the data link layer*



Note that the frame that is travelling between computer A and router R1 may be different from the one travelling between router R1 and R3. When the frame is received by router R1, this router passes the frame to the data link layer protocol shown at the left. The frame is opened, the data are removed. The data are then passed to the data link layer protocol shown at the right to create a new frame to be sent to the router R3. The reason is that the two links, link 1 and link 3, may be using different protocols and require frames of different formats. Note also that the figure does not show the physical movement of frames; the physical movement happens only at the physical layer. The two nodes communicate logically at the data link layer, not physically. In other words, the data link layer at router R1 only thinks that a frame has been sent directly from the data link layer at computer A. What is sent from A to R1 is a stream of bits from one physical layer to another. Since a frame at A is transformed to a stream of bits, and the bits at R1 are transformed to a frame, it gives this impression to the two data link layer that a frame has been exchanged.

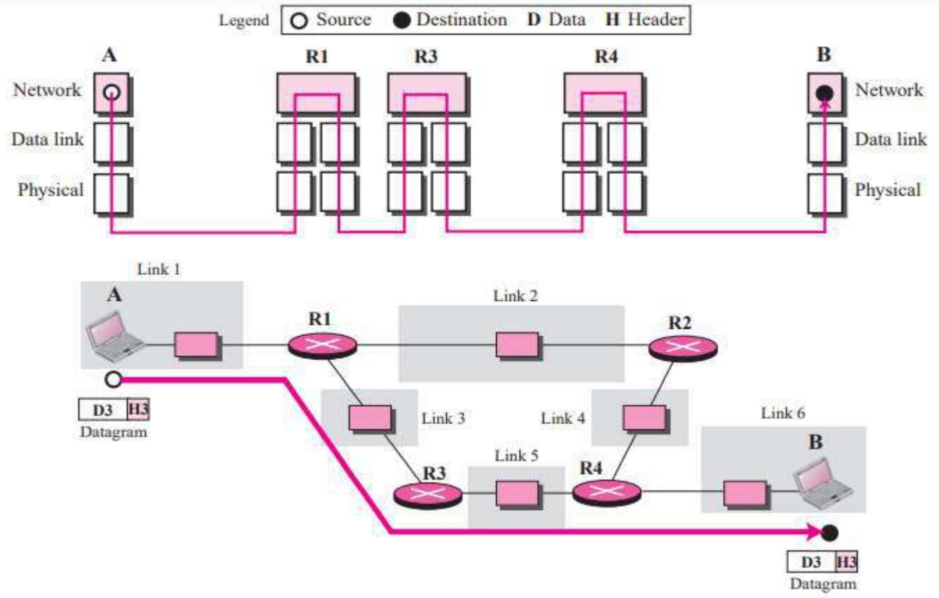
The unit of communication at the data link layer is a frame.

Network Layer

At the network layer (or, more accurately, the internetwork layer), TCP/IP supports the Internet Protocol (IP). The Internet Protocol (IP) is the transmission mechanism used by the TCP/IP protocols. 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.

Figure 2.12 shows the communication at the network layer.

Figure 2.12 *Communication at the network layer*



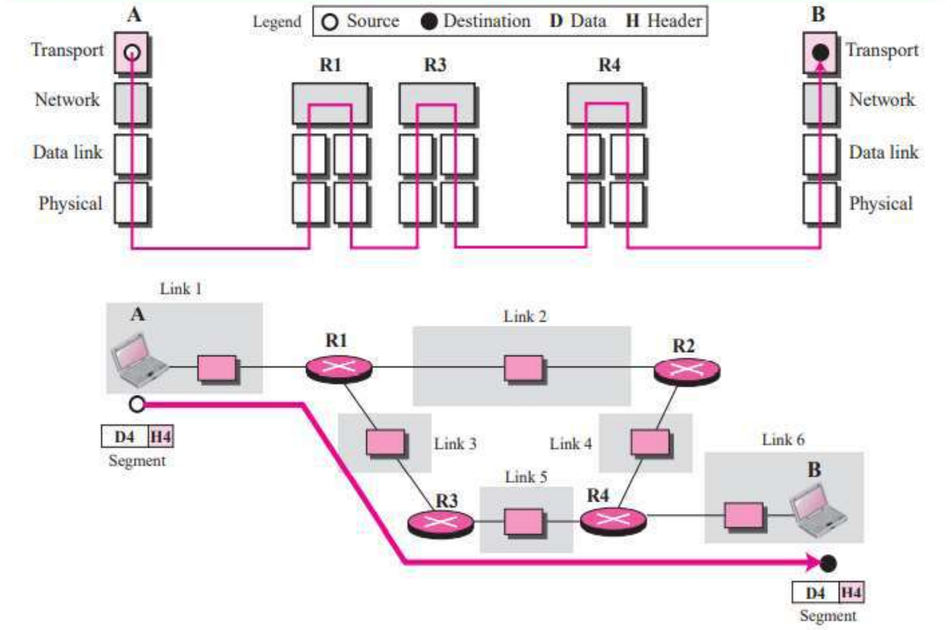
Note that there is a main difference between the communication at the network layer and the communication at data link or physical layers. Communication at the network layer is end to end while the communication at the other two layers are node to node. The datagram started at computer A is the one that reaches computer B. The network layers of the routers can inspect the source and destination of the packet for finding the best route, but they are not allowed to change the contents of the packet. Of course, the communication is logical, not physical. Although the network layer of computer A and B think that they are sending and receiving datagrams, the actual communication again is done at the physical level.

The unit of communication at the network layer is a datagram.

Transport Layer

There is a main difference between the transport layer and the network layer. Although all nodes in a network need to have the network layer, only the two end computers need to have the transport layer. The network layer is responsible for sending individual datagrams from computer A to computer B; the transport layer is responsible for delivering the whole message, which is called a segment, a user data gram, or a packet, from A to B. A segment may consist of a few or tens of datagrams. The segments need to be broken into datagrams and each datagram has to be delivered to the network layer for transmission. Since the Internet defines a different route for each datagram, the datagrams may arrive out of order and may be lost. The trans port layer at computer B needs to wait until all of these datagrams to arrive, assemble them and make a segment out of them. Figure 2.13 shows the communication at the transport layer.

Figure 2.13 *Communication at the transport layer*



Again, we should know that the two transport layers only think that they are communicating with each other using a segment; the communication is done through the physical layer and the exchange of bits.

Traditionally, the transport layer was represented in the TCP/IP suite by two protocols: User Datagram Protocol (UDP) and Transmission Control Protocol (TCP). A new protocol called Stream Control Transmission Protocol (SCTP) has been introduced in the last few years.

The unit of communication at the transport layer is a segment, user datagram, or a packet, depending on the specific protocol used in this layer.

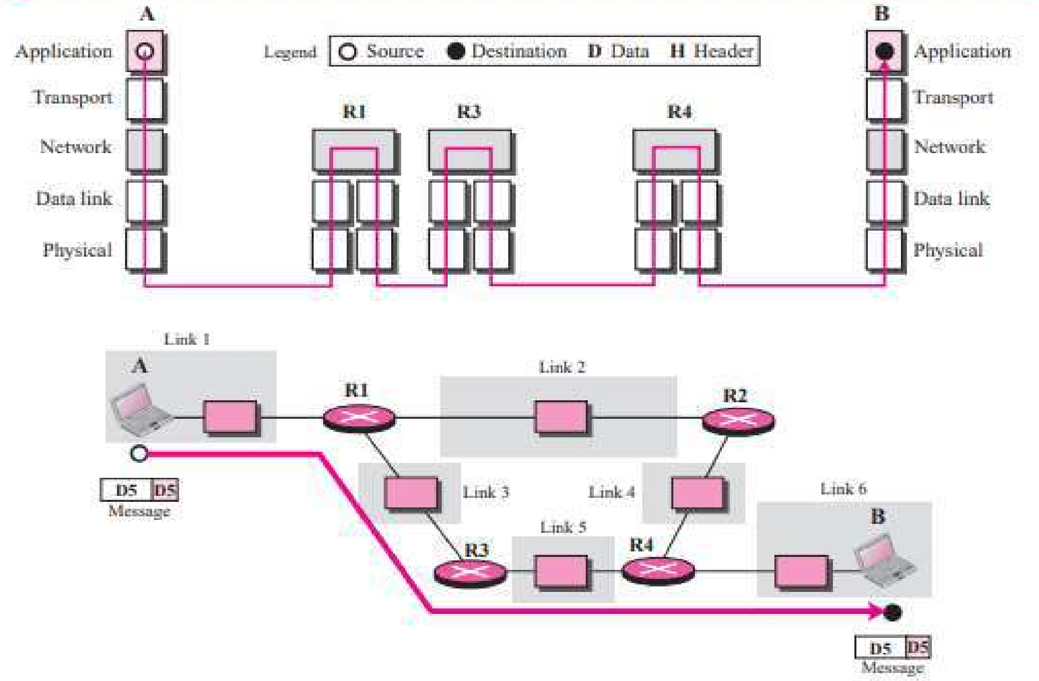
Application Layer

The application layer in TCP/IP is equivalent to the combined session, presentation, and application layers in the OSI model. The application layer allows a user to access the ser vices of our private internet or the global Internet. Many protocols are defined at this layer to provide services such as electronic mail, file transfer, accessing the World Wide Web, and so on. We cover most of the standard protocols in later chapters. Figure 2.14 shows the communication at the application layer.

Note that the communication at the application layer, like the one at the transport layer, is end to end. A message generated at computer A is sent to computer B without being changed during the transmission.

The unit of communication at the application layer is a message.

Figure 2.14 *Communication at the application layer*



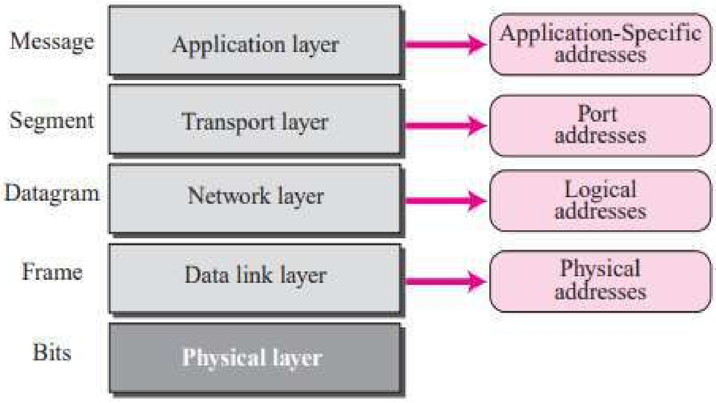
ADDRESSING

Four levels of addresses are used in an internet employing the TCP/IP protocols: physical address, logical address, port address, and application-specific address. Each address is related to a one layer in the TCP/IP architecture, as shown in Figure 2.15.

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 link (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). LocalTalk (Apple), however, has a 1-byte dynamic address that changes each time the station comes up.

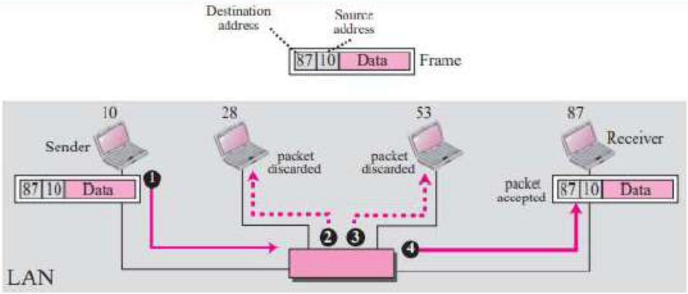
Figure 2.15 *Addresses in the TCP/IP Protocol Suite*



Example 2.3

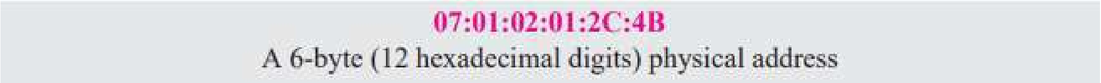
In Figure 2.16 a node with physical address 10 sends a frame to a node with physical address 87. The two nodes are connected by a link (a LAN). At the data link layer, this frame contains physical (link) addresses in the header. These are the only addresses needed. The rest of the header contains other information needed at this level. The trailer usually contains extra bits needed for error detection. As the figure shows, the computer with physical address 10 is the sender, and the computer with physical address 87 is the receiver. The data link layer at the sender receives data from an upper layer. It encapsulates the data in a frame, adding a header and a trailer. The header, among other pieces of information, carries the receiver and the sender physical (link) addresses. Note that in most data link protocols, the destination address 87 in this case, comes before the source address (10 in this case). The frame is propagated through the LAN. Each station with a physical address other than 87 drops the frame because the destination address in the frame does not match its own physical address. The intended destination computer, however, finds a match between the destination address in the frame and its own physical address. The frame is checked, the header and trailer are dropped, and the data part is decapsulated and delivered to the upper layer.

Figure 2.16 *Example 2.3: physical addresses*



Example 2.4

As we will see in Chapter 3, most local area networks use a 48-bit (6-byte) physical address writ ten as 12 hexadecimal digits; every byte (2 hexadecimal digits) is separated by a colon, as shown below:



07:01:02:01:2C:4B

A 6-byte (12 hexadecimal digits) physical address

Unicast, Multicast, and Broadcast Physical Addresses

Physical addresses can be either unicast (one single recipient), multicast (a group of recipients), or broadcast (to be received by all systems in the network). Some networks support all three addresses. For example, Ethernet (see Chapter 3) supports the unicast physical addresses (6 bytes), the multicast addresses, and the broadcast addresses. Some networks do not support the multicast or broadcast physical addresses. If a frame must be sent to a group of recipients or to all systems, the multicast or broadcast address must be simulated using unicast addresses. This means that multiple packets are sent out using unicast addresses.

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.

Example 2.5

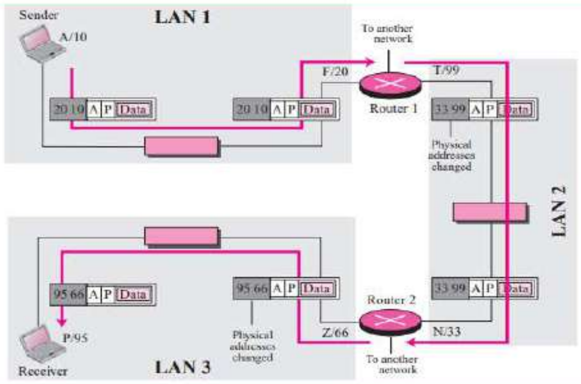
Figure 2.17 shows a part of an internet with two routers connecting three LANs.

Each device (computer or router) has a pair of addresses (logical and physical) for each connection. In this case, each computer is connected to only one link and therefore has only one pair of addresses. Each router, however, is connected to three networks (only two are shown in the figure). So each router has three pairs of addresses, one for each connection. Although it may be obvious that each router must have a separate physical address for each connection, it may not be obvious why it needs a logical address for each connection.

The computer with logical address A and physical address 10 needs to send a packet to the computer with logical address P and physical address 95. We use letters to show the logical addresses and numbers for physical addresses, but note that both are actually numbers, as we will see in later chapters.

The sender encapsulates its data in a packet at the network layer and adds two logical addresses (A and P). Note that in most protocols, the logical source address comes before the logical destination address (contrary to the order of physical addresses). The network layer, however, needs to find the physical address of the next hop before the packet can be delivered. The network layer consults its routing table (see Chapter 6) and finds the logical address of the next hop (router 1) to be F. Another protocol, Address Resolution Protocol (ARP), which will be discussed in later chapters, finds the physical address of router 1 that corresponds to its logical address (20). Now the network layer passes this address to the data link layer, which in turn, encapsulates the packet with physical destination address 20 and physical source address 10.

Figure 2.17 *Example 2.5: logical addresses*



The frame is received by every device on LAN 1, but is discarded by all except router 1, which finds that the destination physical address in the frame matches with its own physical address. The router decapsulates the packet from the frame to read the logical destination address P. Since the logical destination address does not match the router's logical address, the router knows that the packet needs to be forwarded. The router consults its routing table and ARP to find the physical destination address of the next hop (router 2), creates a new frame, encapsulates the packet, and sends it to router 2.

Note the physical addresses in the frame. The source physical address changes from 10 to 99. The destination physical address changes from 20 (router 1 physical address) to 33 (router 2 physical address). The logical source and destination addresses must remain the same; otherwise the packet will be lost.

At router 2 we have a similar scenario. The physical addresses are changed, and a new frame is sent to the destination computer. When the frame reaches the destination, the packet is decapsulated. The destination logical address P matches the logical address of the computer. The data are decapsulated from the packet and delivered to the upper layer. Note that although physical addresses will change from hop to hop, logical addresses remain the same from the source to destination.

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

Unicast, Multicast, and Broadcast Addresses

The logical addresses can be either unicast (one single recipient), multicast (a group of recipients), or broadcast (all systems in the network). There are limitations on broadcast addresses.

Port Addresses

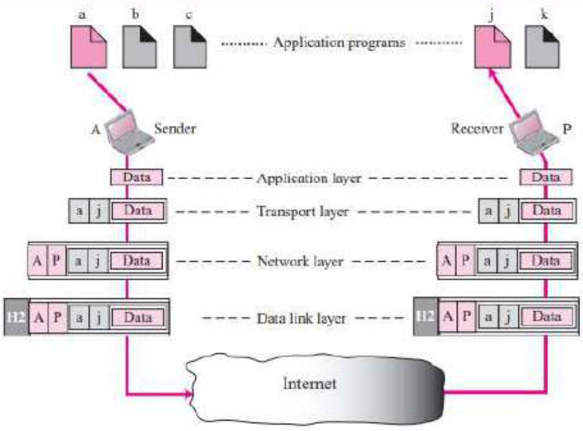
The IP address and the physical address are necessary for a quantity of data to travel from a source to the destination host. However, arrival at the destination host is not the final objective of data communications on the Internet. A system that sends nothing but data from one computer to another is 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 TCP/IP architecture, the label assigned to a process is called a port address. A port address in TCP/IP is 16 bits in length.

Example 2.6

Figure 2.18 shows two computers communicating via the Internet. The sending computer is running three processes at this time with port addresses a, b, and c. The receiving computer is running two processes at this time with port addresses j and k. Process a in the sending computer needs to communicate with process j in the receiving computer. Note that although both computers are using the same application, FTP, for example, the port addresses are different because one is a client program and the other is a server program, as we will see in Chapter 17. To show that data from process a need to be delivered to process j, and not k, the transport layer encapsulates data from the application layer in a packet and adds two port addresses (a and j), source and destination. The packet from the transport layer is then encapsulated in another packet at the network layer with logical source and destination addresses (A and P). Finally, this packet is encapsulated in a frame with the physical source and destination addresses of the next hop. We have not shown the physical addresses because they change from hop to hop inside the cloud designated as the Internet. Note that although physical addresses change from hop to hop, logical and port addresses remain the same from the source to destination. There are some exceptions to this rule that we discuss later in the book.

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

Figure 2.18 *Example 2.6: port numbers*



Example 2.7

As we will see in future chapters, a port address is a 16-bit address represented by one decimal number as shown.

753

A 16-bit port address represented as one single number

Application-Specific Addresses

Some applications have user-friendly addresses that are designed for that specific appli cation. 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, as we will see in later chapters.