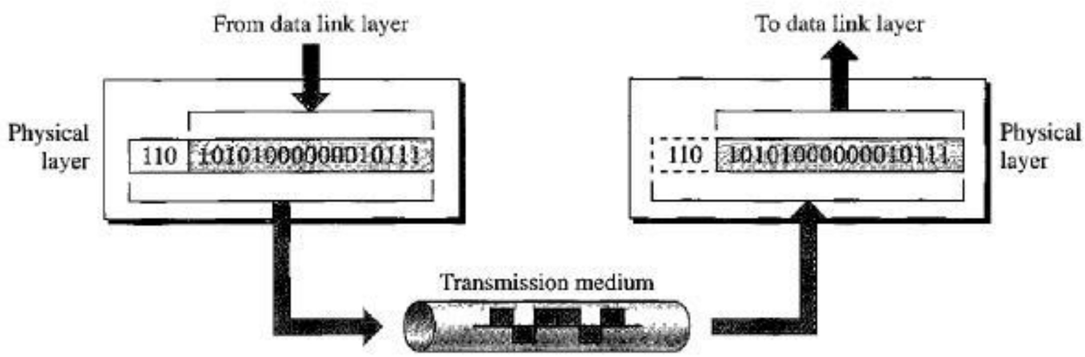
Unit 2: The Physical Layer

Functions of Physical Layer:

The physical layer coordinates the functions required to carry a bit stream over a physical medium. It deals with the mechanical and electrical specifications of the interface and transmission medium. It also defines the procedures and functions that physical devices and interfaces have to perform for transmission to occur. Figure below shows the position of the physical layer with respect to the transmission medium and the data link layer.

Figure 2.5 *Physical layer*



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

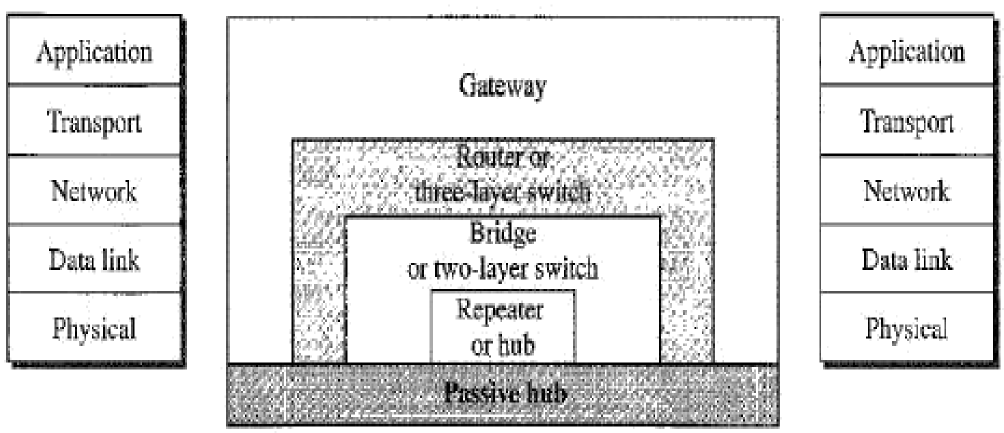
Specific responsibilities of physical layer include:

- Physical characteristics of interfaces and medium
- Representation of bits
- Data rate
- Synchronization of bits
- Line configuration
- Transmission mode

Network Devices (Connecting Devices):

We divide connecting devices into five different categories based on the layer in which they operate in a network, as shown in Figure below.

Figure 15.1 *Five categories of connecting devices*



Data and Signals:

Signal:

A signal is an electrical or electromagnetic current that is used for carrying data from one device or network to another. It is the key component behind data communication and networking. Signals can be periodic and no periodic. A periodic signal repeats the pattern over identical periods. A no periodic signal changes without repeating a pattern or cycle over time.

A signal can be either analog or digital.

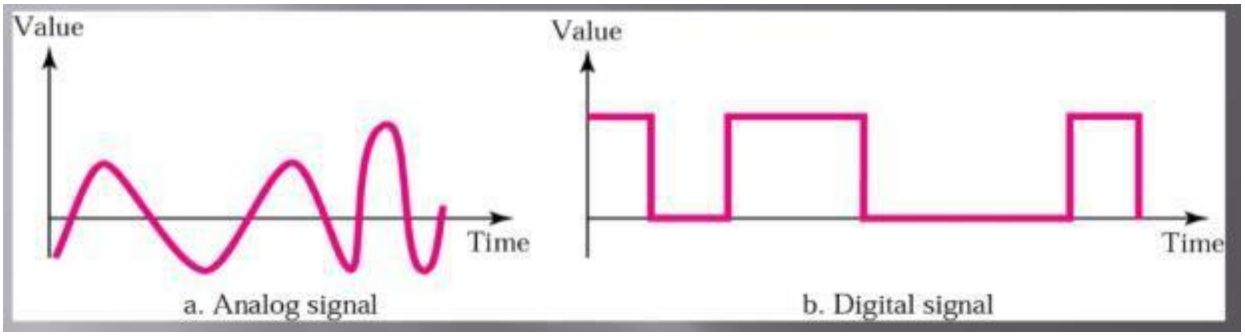
Analog Signal:

Analog signal is a continuous wave that keeps on changing over a time period. In other words, an analog signal is a continuous wave denoted by a sine wave and may vary in signal strength (amplitude) or frequency (waves per unit time). Analog signals can be classified as simple or composite. A simple analog signal or sine wave cannot be further decomposed into simpler signals. A composite analog signal is composed of multiple sine waves.

Digital Signal:

Digital signals also carry information like analog signals but is somewhat is different from analog signals. Digital signal is noncontiguous, discrete time signal. Digital signal carries information or data in the binary.

Form i.e. a digital signal represent information in the form of bits (0s and 1s). Digital signals are easier to transmit and are more reliable when compared to analog signals.



Key Differences Between Analog and Digital Signal:

- An analog signal represents a continuous wave that keeps changing over a time period. On the other hand, a digital signal represents a noncontiguous wave that carries information in a binary format and has discrete values.
- An analog signal is always represented by the continuous sine wave whereas, a digital signal is represented by square waves.
- While talking of analog signal, we describe the behavior of the wave in respect of amplitude, period or frequency, and phase of the wave. On the other hand, while talking of discrete signals we describe the behavior of the wave in respect of bit rate and bit interval.
- The range of an analog signal is not fixed whereas the range of the digital signal is finite and which can be 0 or 1.
- An analog signal is more prone to distortion in response to noise, but a digital signal has immunity in response to noise hence it rarely faces any distortion.
- An analog signal transmits data in the form of wave whereas, a digital signal transmits the data in the binary form i.e. in the form of bits.

- The best example of an analog signal is a human voice, and the best example of a digital signal is the transmission of data in a computer.

Analog Data:

The term analog data refers to information that is continuous; For example, an analog clock that has hour, minute, and second hands gives information in a continuous form; the movements of the hands are continuous. Analog data, such as the sounds made by a human voice, take on continuous values. When someone speaks, an analog wave is created in the air. This can be captured by a microphone and converted to an analog signal or sampled and converted to a digital signal.

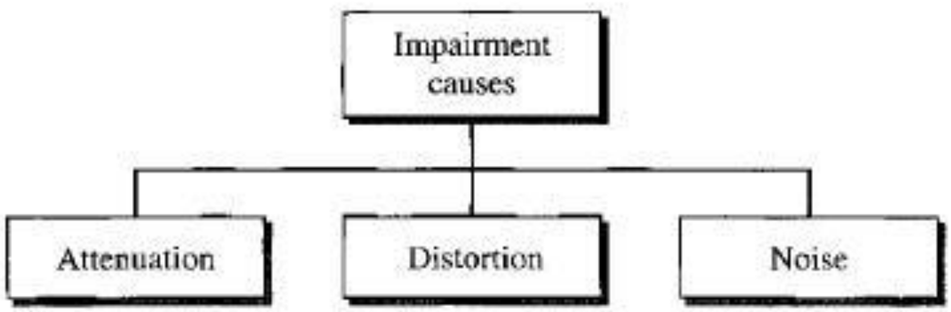
Digital Data:

Digital data refers to information that has discrete states. For example, a digital clock that reports the hours and the minutes will change suddenly from 8:05 to 8:06. Digital data takes on discrete values. For example, data are stored in computer memory in the form of 0s and 1s. They can be converted to a digital signal or modulated into an analog signal for transmission across a medium.

Transmission Impairment:

Signals travel through transmission media, which are not perfect. The imperfection causes signal impairment. This means that the signal at the beginning of the medium is not the same as the signal at the end of the medium. What is sent is not what is received. Three causes of impairment are attenuation, distortion, and noise.

Causes of impairment



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 resistance of the medium. That is why a wire carrying electric signals gets warm, if not hot, after a while. Some of the electrical energy in the signal is converted to heat. To compensate for this loss, amplifiers are used to amplify the signal. Attenuation is measured in terms of Decibels.

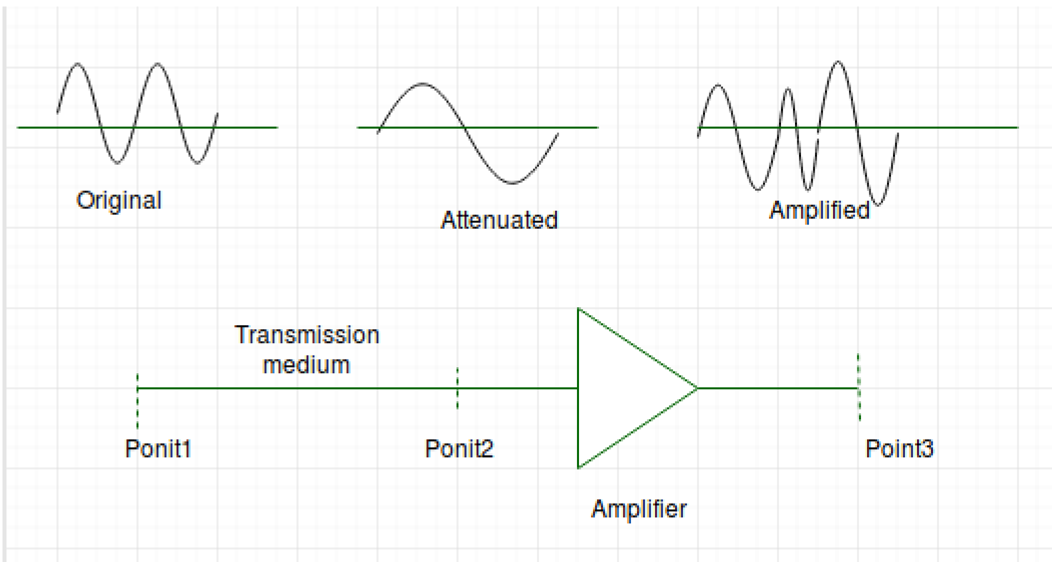
The decibel (dB) measures the relative strengths of two signals or one signal at two different points. Note that the decibel is negative if a signal is attenuated and positive if a signal is amplified.

$dB = 10\log_{10} P2/P1$

Variables P1 and P2 are the powers of a signal at points 1 and 2, respectively.

or equivalently - when comparing two signals with the same kind of waveform

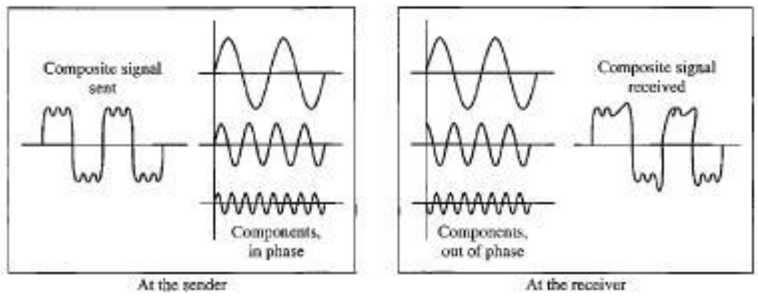
$dB = 20\log_{10} P2/P1$



Distortion:

Distortion means that the signal changes its form or shape. Distortion 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 final destination. Differences in delay may create a difference in phase if the delay is not exactly the same as the period duration. In other words, signal components at the receiver have phases different from what they had at the sender. The shape of the composite signal is therefore not the same. Figure shows the effect of distortion on a composite signal.

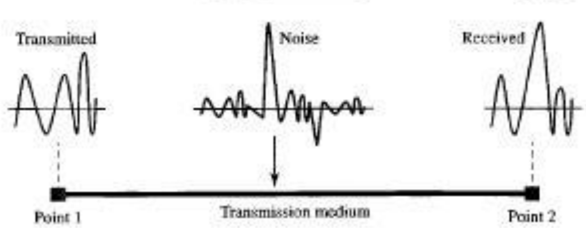
Figure 3.28 *Distortion*



Noise

Noise is another cause of impairment. Several types of noise, such as thermal noise, induced noise, crosstalk, and impulse noise, may corrupt the signal. Thermal noise is the random motion of electrons in a wire which creates an extra signal not originally sent by the transmitter. Induced noise comes from sources such as motors and appliances. These devices act as a sending antenna, and the transmission medium acts as the receiving antenna. Crosstalk is the effect of one wire on the other. One wire act as a sending antenna and the other as the receiving antenna. Impulse noise is a spike (a signal with high energy in a very short time) that comes from power lines, lightning, and so on.

Figure 3.29 *Noise*



Signal-to-Noise Ratio (SNR)

The signal-to-noise ratio is defined as

SNR= Average Signal power / Average Noise Power

SNR is actually the ratio of what is wanted (signal) to what is not wanted (noise). A high SNR means the signal is less corrupted by noise; a low SNR means the signal is more corrupted by noise. Because SNR is the ratio of two powers, it is often described in decibel units, SNR dB, defined as

$$SNR_{dB} = 10\log_{10} SNR$$

Data Rate Limits

Data rate can be defined as how fast can we send the data, in bits per second, over a channel.

Data rate depends on three factors:

- The bandwidth available
- The level of the signals we use
- The quality of the channel (the level of noise)

There are two theoretical formulas to calculate the data rate:

- Nyquist for a noiseless channel
- Shannon for a noisy channel

Noiseless Channel: Nyquist Bit Rate

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

$$BitRate = 2 * Bandwidth * \log_2(L)$$

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.

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 * $\log_2$(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.

Performance

One important issue in networking is the performance of the network-how good it is? It can be referred as Quality of Service (QoS) that is an overall measurement of the network performance. There are four factors of determining network performance.

- Bandwidth
- Throughput
- Delay
- Jitter

Bandwidth

One characteristic that measures network performance is bandwidth. However, the term can be used in two different contexts with two different measuring values: bandwidth in hertz and bandwidth in bits per second.

- The first, bandwidth in hertz, refers to the range of frequencies in signal or the range of frequencies that a channel can pass. For example, we can say the bandwidth of a subscriber telephone line is 4 kHz.
- The second, bandwidth in bits per second, refers to the speed of bit transmission in a channel or link. For example, one can say the bandwidth of a Fast Ethernet network (or the links in this network) is a maximum of 100 Mbps. This means that this network can send 100 Mbps.

Throughput

- The throughput is a measure of how fast we can actually send data through a network.

- For example, we may have a link with a bandwidth of 1 Mbps, but the devices connected to the end of the link may handle only 200 kbps. This means that we cannot send more than 200 kbps through this link.

Delay (Latency)

- The latency or delay defines 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.
- We can say that latency is made of four components: propagation time, transmission time, queuing time and processing delay.

  Latency =propagation time +transmission time +queuing time + processing delay

- Propagation time measures the time required for a bit to travel from the source to the destination. The propagation time is calculated by dividing the distance by the propagation speed.

Propagation Time= Distance / Propagation Speed

The time required for transmission of a message depends on the size of the message and the bandwidth of the channel.

Transmission Time= Message Size / Bandwidth

- The third component in latency is the queuing time, the time needed for each intermediate or end device to hold the message before it can be processed. The queuing time is not a fixed factor; it changes with the load imposed on the network. When there is heavy traffic on the network, the queuing time increases. An intermediate device, such as a router, queues the arrived messages and processes them one by one. If there are many messages, each message will have to wait.

Jitter

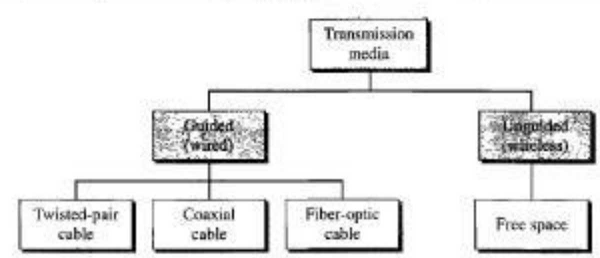
- Another performance issue that is related to delay is jitter. We can roughly say that jitter is a problem if different packets of data encounter different delays and the application using the data at the receiver site is time-sensitive (audio and video data, for example).
- For Example, If the delay for the first packet is 20 ms, for the second is 45 ms, and for the third is 40 ms, then the real-time application that uses the packets endures jitter.

TRANSMISSION MEDIA:

A transmission medium can be broadly defined as anything that can carry information from a source to a destination.

Classification of Transmission Media:

Figure 7.2 *Classes of transmission media*



Guided Media

Guided media, which are those that provide a channel from one device to another, include twisted-pair cable, coaxial cable, and fiber-optic cable. A signal traveling along any of these media is directed and contained by the physical limits of the medium. Twisted-pair and coaxial cable use metallic (copper) conductors that accept and transport signals in the form of electric current. Optical fiber is a cable that accepts and transports signals in the form of light.

1. Twisted-Pair Cable

A twisted pair consists of two conductors (normally copper), each with its own plastic insulation, twisted together, as shown in Figure below.

Figure 7.3 *Twisted-pair cable*

One of the wires is used to carry signals to the receiver, and the other is used only as a ground reference.

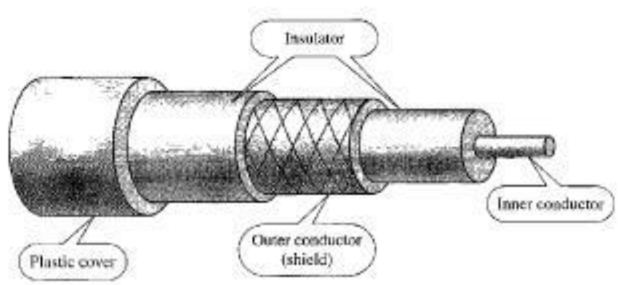
Applications

Twisted-pair cables are used in telephone lines to provide voice and data channels. The DSL lines that are used by the telephone companies to provide high-data-rate connections also use the high-bandwidth capability of unshielded twisted-pair cables. Local-area networks also use twisted-pair cables.

2. Coaxial Cable

Coaxial cable (or coax) carries signals of higher frequency ranges than those in twisted pair cable, in part because the two media are constructed quite differently. The outer metallic wrapping serves both as a shield against noise and as the second conductor, which completes the circuit. This outer conductor is also enclosed in an insulating sheath, and the whole cable is protected by a plastic cover (see Figure below).

Figure 7.7 *Coaxial cable*



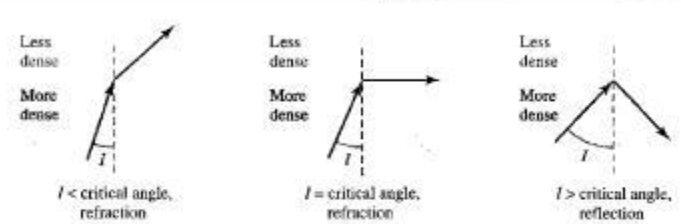
Applications

Cable TV networks also use coaxial cables. In the traditional cable TV network, the entire network used coaxial cable. Later, however, cable TV providers replaced most of the media with fiber-optic cable;

3. Fiber Optic Cable:

A fiber-optic cable is made of glass or plastic and transmits signals in the form of light. To understand optical fiber, we first need to explore several aspects of the nature of light. Light travels in a straight line as long as it is moving through a single uniform If a ray of light traveling through one substance suddenly enters another substance (of a different density), the ray changes direction. Figure below shows how a ray of light changes direction when going from a denser to a less dense substance.

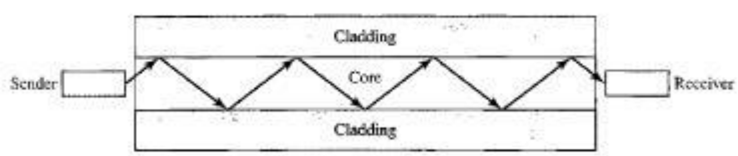
Figure 7.10 *Bending of light ray*



As the figure shows, if the angle of incidence I (the angle the ray makes with the line perpendicular to the interface between the two substances) is less than the critical angle, the ray refracts 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 (makes a turn) and travels again in the denser substance. Note that the critical angle is a property of the substance, and its value differs from one substance to another.

Optical fibers use reflection to guide light through a channel. A glass or plastic core is surrounded by a cladding of less dense glass or plastic. The difference in 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. See Figure below.

Figure 7.11 *Optical fiber*



<u>Applications</u>

Fiber-optic cable is often found in backbone networks because its wide bandwidth is cost-effective. We can transfer data at a rate of 1600 Gbps.

Advantages and Disadvantages of Optical Fiber

<u>Advantages</u>

Fiber-optic cable has several advantages over metallic cable (twisted pair or coaxial).

a. <u>Higher bandwidth</u>. Fiber-optic cable can support dramatically higher bandwidths (and hence data rates) than either twisted-pair or coaxial cable. Currently, data rates and bandwidth utilization over fiber-optic cable are limited not by the medium but by the signal generation and reception technology available.

b. Less signal attenuation. Fiber-optic transmission distance is significantly greater than that of other guided media. A signal can run for 50 km without requiring regeneration. We need repeaters every 5 km for coaxial or twisted-pair cable.

c. Immunity to electromagnetic interference. Electromagnetic noise cannot affect fiber-optic cables.

d. Light weight. Fiber-optic cables are much lighter than copper cables.

e. Greater immunity to tapping. Fiber-optic cables are more immune to tapping than copper cables. Copper cables create antenna effects that can easily be tapped.

Disadvantages

There are some disadvantages in the use of optical fiber.

a. Installation and maintenance. Fiber-optic cable is a relatively new technology. Its installation and maintenance require expertise that is not yet available everywhere.

b. Cost. The cable and the interfaces are relatively more expensive than those of other guided media. If the demand for bandwidth is not high, often the use of optical fiber cannot be justified.

UNGUIDED MEDIA: WIRELESS

Unguided media transport electromagnetic waves without using a physical conductor. This type of communication is often referred to as wireless communication. Signals are normally broadcast through free space and thus are available to anyone who has a device capable of receiving them.

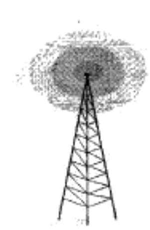
1. Radio Waves

Waves ranging in frequencies between 3 kHz and 1 GHz are called radio w aves. Radio waves, for the most part, are omnidirectional. When an antenna transmits radio waves, they are propagated in all directions.

This means that the sending and receiving antennas do not have to be aligned. A sending antenna sends waves that can be received by any receiving antenna.

The omnidirectional property has a disadvantage, too. The radio waves transmitted by one antenna are susceptible to interference by another antenna that may send signals using the same frequency or band. This makes radio waves a good candidate for long-distance broadcasting. Advantage → a radio waves can receive signals inside a building.

Figure 7.20 *Omnidirectional antenna*



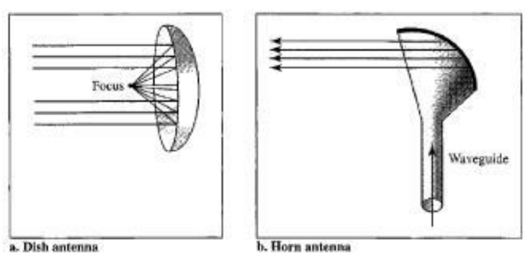
Applications: The omnidirectional characteristics of radio waves make them useful for multicasting, in which there is one sender but many receivers. FM radio, television, maritime radio, cordless phones are examples of multicasting.

2. Microwaves

Electromagnetic waves having frequencies between 1 and 300 GHz are called microwaves. Microwaves are unidirectional. When an antenna transmits microwave waves, they can be narrowly focused.

This means that the sending and receiving antennas need to be aligned. The unidirectional property has an obvious advantage. A pair of antennas can be aligned without interfering with another pair of aligned antennas.

Figure 7.21 *Unidirectional antennas*



Microwaves need unidirectional antennas that send out signals in one direction. Two types of antennas are used for microwave communications: the parabolic dish and the horn (see Figure).

3. Infrared

Infrared waves, with frequencies from 300 GHz to 400 THz, can be used for short-range communication. Infrared waves, having high frequencies, cannot penetrate walls.

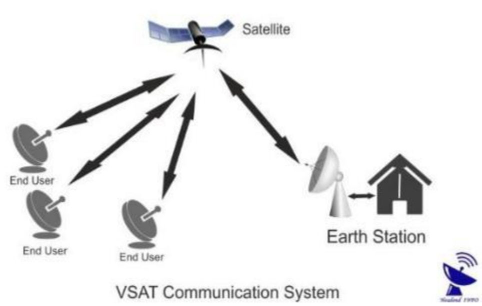
This advantageous characteristic prevents interference between one system and another; a short-range communication system in one room cannot be affected by another system in the next room.

When we use our infrared remote control, we do not interfere with the use of the remote by our neighbors. However, this same characteristic makes infrared signals useless for long-range communication. In addition, we cannot use infrared waves outside a building because the sun's rays contain infrared waves that can interfere with the communication.

VSAT:

A very small aperture terminal (VSAT) is a small telecommunication earth station that receives and transmits real-time data via satellite. VSAT is a satellite communications system that serves home and business users.

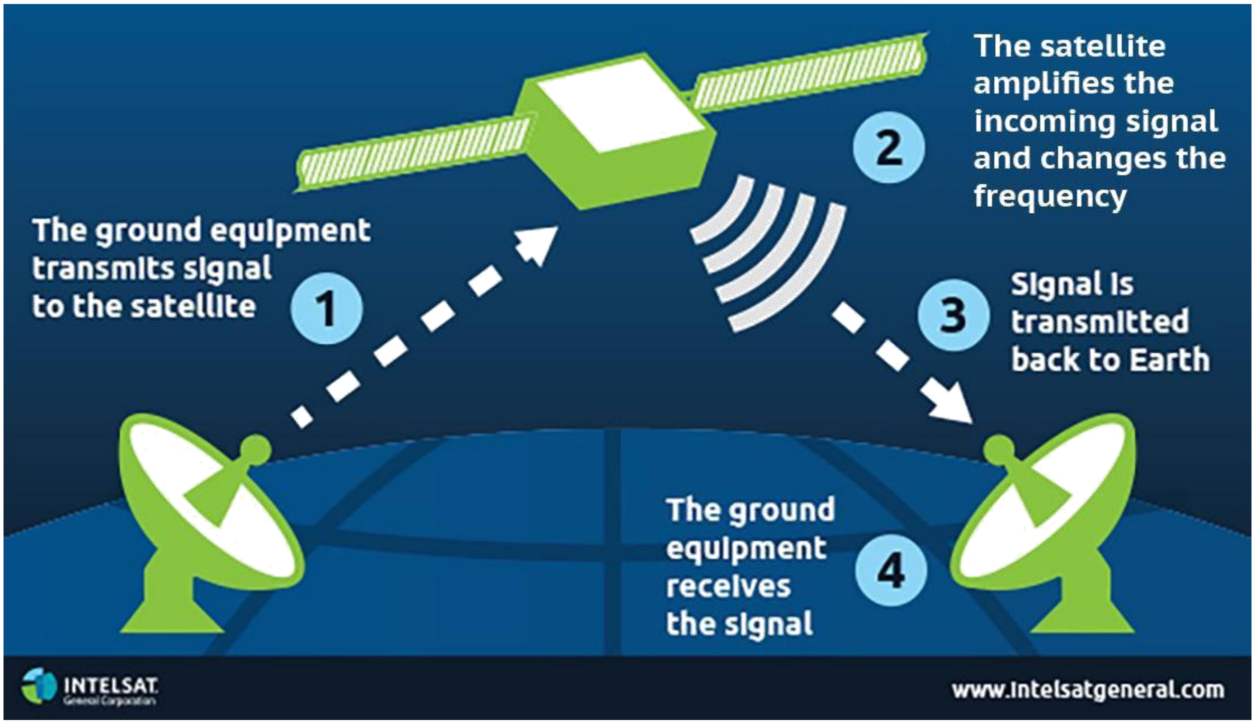
The satellite sends and receives signals from an earth station computer that acts as a hub for the system. For one end user to communicate with another, each transmission has to first go to the hub station which retransmits it via the satellite to the other end user's VSAT.



A main advantage of VSAT is it provides companies with complete control over their own communications infrastructure without having to depend upon third party sources.

<u>Satellite:</u>

A communications satellite is an artificial satellite that relays and amplifies radio telecommunications signals via a transponder; it creates a communication channel between a source transmitter and a receiver at different locations on Earth. Communications satellites are used for television, telephone, radio, internet, and military applications. The purpose of communications satellites is to relay the signal around the curve of the Earth allowing communication between widely separated geographical points.



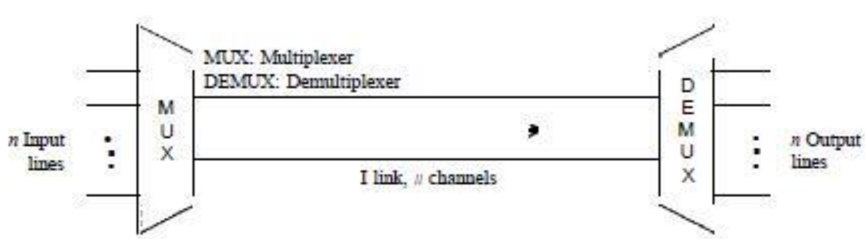
Applications: Television, Internet, Military

Bandwidth Utilization: Multiplexing and Spreading

- There are two broad categories of bandwidth utilization: multiplexing and spreading.

Multiplexing

- Multiplexing is the set of techniques that allows the simultaneous transmission of multiple signals across a single data link.
- In a multiplexed system, n lines share the bandwidth of one link.
- MUX combines the streams into a single stream at the sender side whereas the DEMUX separates the streams back into its component transmissions.



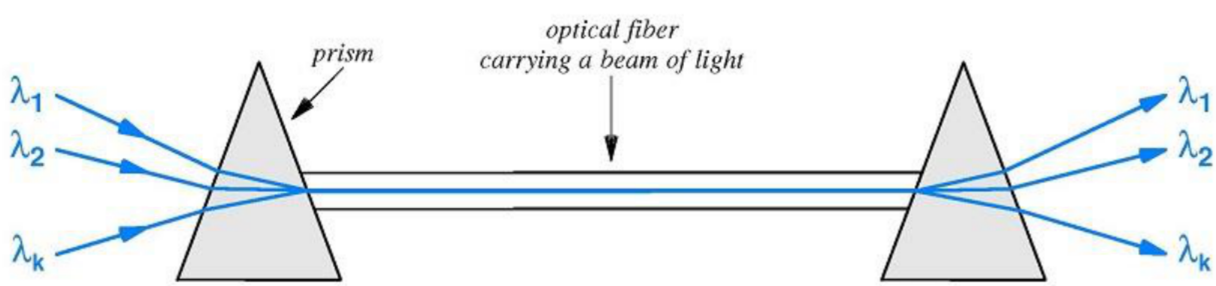
- There are three basic multiplexing techniques: frequency-division multiplexing, wavelength-division multiplexing, and time-division multiplexing.
- The first two are techniques designed for analog signals, the third, for digital signals.

Frequency-Division Multiplexing

- Frequency-division multiplexing (FDM) is an analog technique that can be applied when the bandwidth of a link (in hertz) is greater than the combined bandwidths of the signals to be transmitted.
- In FDM, signals generated by each sending device modulate different carrier frequencies. These modulated signals are then combined into a single composite signal that can be transported by the link.

Wavelength-Division Multiplexing

- Wavelength-division multiplexing (WDM) is designed to use the high-data-rate capability of fiber-optic cable.
- Multiplexing allows us to combine several lines into one.
- WDM is conceptually the same as FDM, except that the multiplexing and demultiplexing involve optical signals transmitted through fiber-optic channels.



- Ex- A prism bends a beam of light based on the angle of incidence and the frequency.

Time-Division Multiplexing (TDM)

- TDM is a digital process that allows several connections to share the high bandwidth of a line.
- Each connection occupies a portion of time in the link.
- Note that the same link is used as in FDM; here, however, the link is shown sectioned by time rather than by frequency.

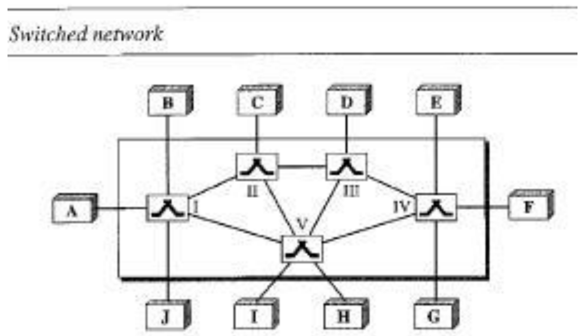
Spread Spectrum:

- In spread spectrum, we also combine signals from different sources to fit into a larger bandwidth, but our goals are somewhat different.
- Spread spectrum is designed to be used in wireless applications (LANs and WANs).
- The expanded bandwidth allows the source to wrap its message in a protective envelope for a more secure transmission.

Switching:

A network is a set of connected devices. Whenever we have multiple devices, we have the problem of how to connect them to make one-to-one communication possible. One solution is to make a point-to-point connection between each pair of devices (Mesh topology) or between a central device and every other device (a star topology). These methods are not applicable for very large networks.

A better solution is **switching**. A switched network consists of a series of interlinked nodes, called switches. Switches are devices capable of creating temporary connections between two or more devices linked to the switch.



We can divide today's network into three broad categories: circuit-switched networks, packet-switched networks and message-switched networks.

Circuit Switched Networks:

A circuit-switched network consists of a set of switches connected by physical links. A connection between two stations is a dedicated path made of one or more links. However, each connection uses only one dedicated channel on each link.

The end systems, such as computers or telephones, are directly connected to a switch. We have shown only two end systems for simplicity. When end system A needs to communicate with end system M, system A needs to request a connection to M that must be accepted by all switches as well as by M itself.

Three Phases:

The actual communication in a circuit-switched network requires three phases: connection setup, data transfer, and connection teardown.

Setup Phase

Before the two parties (or multiple parties in a conference call) can communicate, a dedicated circuit (combination of channels in links) needs to be established.

Data Transfer Phase

After the establishment of the dedicated circuit (channels), the two parties can transfer data.

Teardown Phase

When one of the parties needs to disconnect, a signal is sent to each switch to release the resources.

Packet Switched Networks:

A packet switched network (PSN) is a type of computer communications network that groups and sends data in the form of small packets. It enables the sending of data or network packets between a source and destination node over a network channel that is shared between multiple users and/or applications.

Packet switching may be classified into connectionless packet switching, also known as datagram switching, and connection-oriented packet switching, also known as virtual circuit switching.

Datagram Approach:

In a datagram network, each packet is treated independently of all others. Even if a packet is part of a multipacket transmission, the network treats it as though it existed alone.

Virtual Circuit Networks:

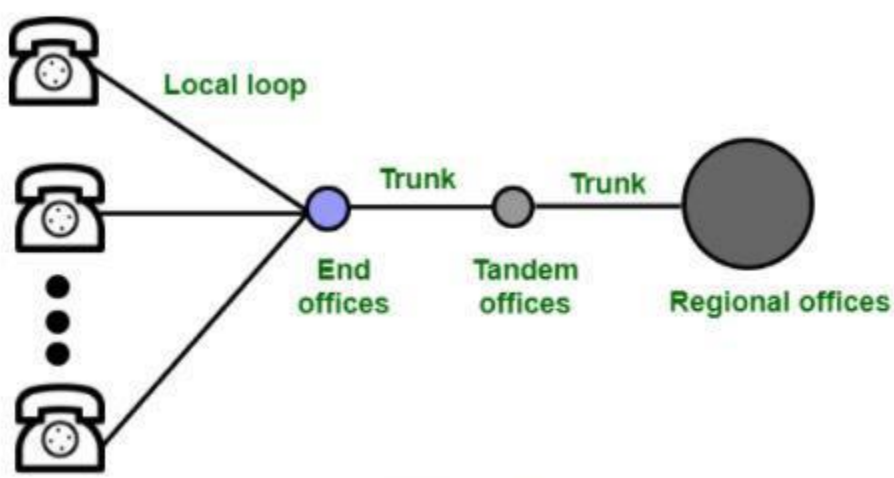
A virtual circuit network is a cross between a circuit-switched network and a datagram network. It has some characteristics of both.

- A virtual-circuit network is normally implemented in the data link layer, while a circuit-switched network is implemented in the physical layer and a datagram network is implemented in the network layer.

Telephone Network:

- Telephone Network is used to provide voice communication which uses Circuit Switching.

  The telephone network is made of three major components: local loops, trunks, and switching offices.



A telephone system

Local Loops

- One component of the telephone network is the local loop, a twisted-pair cable that connects the subscriber telephone to the nearest end office or local central office.

Trunks

- Trunks are transmission media that handle the communication between offices.
- A trunk normally handles hundreds or thousands of connections through multiplexing.

Switching Offices

- To avoid having a permanent physical link between any two subscribers, the telephone company has switches located in a switching office.
- A switch connects several local loops or trunks and allows a connection between different subscribers.

Mobile Networks:

- Mobile Networks or Cellular networks are high-speed, high-capacity voice and data communication networks with enhanced multimedia and seamless roaming capabilities for supporting cellular devices (wireless end devices).
- With the increase in popularity of cellular devices, these networks are used for more than just entertainment and phone calls.
- Cellular telephony is designed to provide communications between two moving units, called mobile stations (MSs), or between one mobile unit and one stationary unit, often called a land unit.

Cable Networks:

- The cable TV network started as a video service provider, but it has moved to the business of Internet access.

<u>Traditional Cable Networks:</u>

- Cable TV started to distribute broadcast video signals to locations with poor or no reception in the late 1940s.
- It was called community antenna s) because an antenna at the top of a tall hill or building received the signals from the TV stations and distributed them, via coaxial cables, to the community.