

COMPUTER NETWORKS

Lecture Notes

DEPARTMENT OF ECE

LENDI INSTITUTE OF ENGINEERING & TECHNOLOGY

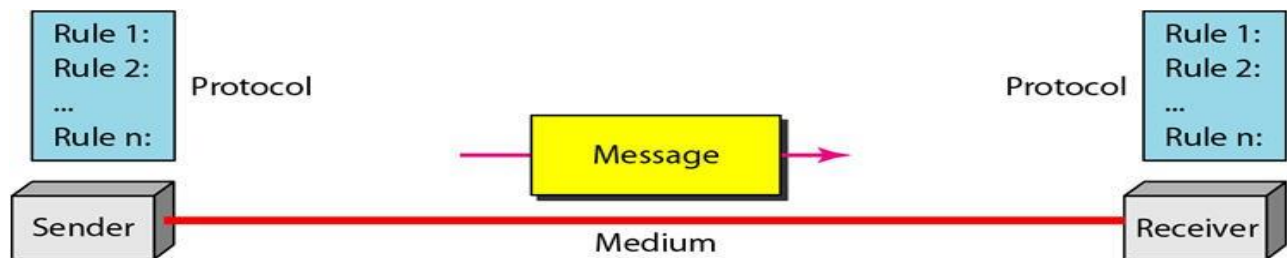
UNIT -I

Introduction to Computer Networks

1.1 Data Communication: When we communicate, we are sharing information. This sharing can be local or remote. Between individuals, local communication usually occurs face to face, while remote communication takes place over distance.

1.1.1 Components:

A data communications system has five components.



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

1.1.2 Data Representation:

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

Text:

In data communications, text is represented as a bit pattern, a sequence of bits (Os or Is). Different sets of bit patterns have been designed to represent text symbols. Each set is called a code, and the process of representing symbols is called coding. Today, the prevalent coding system is called Unicode, which uses 32 bits to represent a symbol or character used in any language in the world. The American Standard Code for Information Interchange (ASCII), developed some decades ago in the United States, now constitutes the first 127 characters in Unicode and is also referred to as Basic Latin.

Numbers:

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

Images:

Images are also represented by bit patterns. In its simplest form, an image is composed of a matrix of pixels (picture elements), where each pixel is a small dot. The size of the pixel depends on the *resolution*. For example, an image can be divided into 1000 pixels or 10,000 pixels. In the second case, there is a better representation of the image (better resolution), but more memory is needed to store the image. After an image is divided into pixels, each pixel is assigned a bit pattern. The size and the value of the pattern depend on the image. For an image made of only blackand-white dots (e.g., a chessboard), a 1-bit pattern is enough to represent a pixel. If an image is not made of pure white and pure black pixels, you can increase the size of the bit pattern to include gray scale. For example, to show four levels of gray scale, you can use 2-bit patterns. A black pixel can be represented by 00, a dark gray pixel by 01, a light gray pixel by 10, and a white pixel by 11. There are several methods to represent color images. One method is called RGB, so called because each color is made of a combination of three primary colors: *red*, green, and blue. The intensity of each color is measured, and a bit pattern is assigned to it. Another method is called YCM, in which a color is made of a combination of three other primary colors: yellow, cyan, and magenta.

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

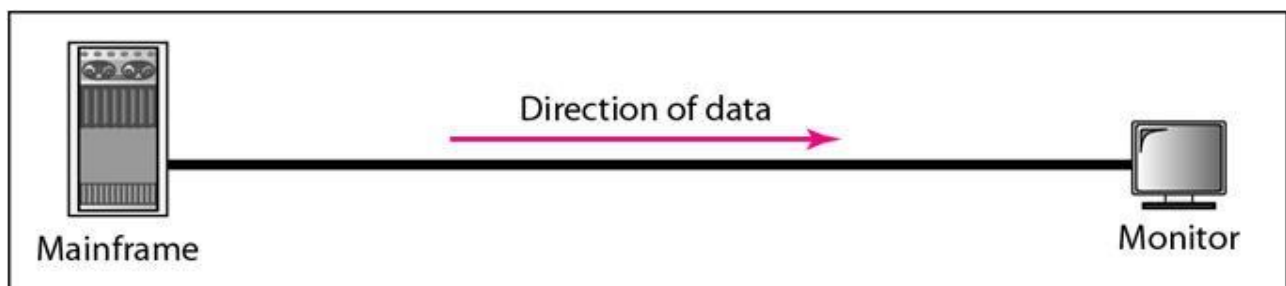
learn how to change sound or music to a digital or an analog signal.

Video:

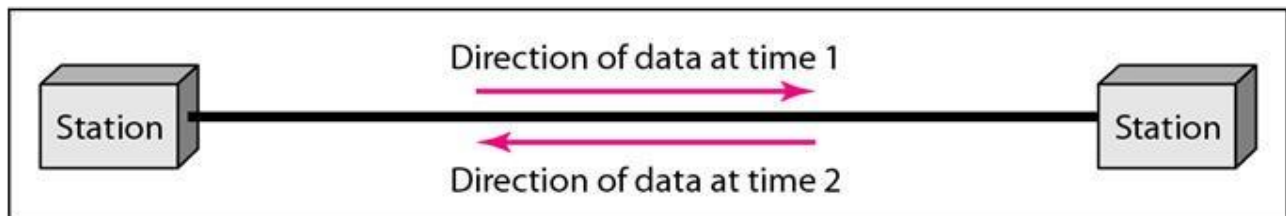
Video refers to the recording or broadcasting of a picture or movie. Video can either be produced as a continuous entity (e.g., by a TV camera), or it can be a combination of images, each a discrete entity, arranged to convey the idea of motion. Again we can change video to a digital or an analog signal.

1.1.3 Data Flow

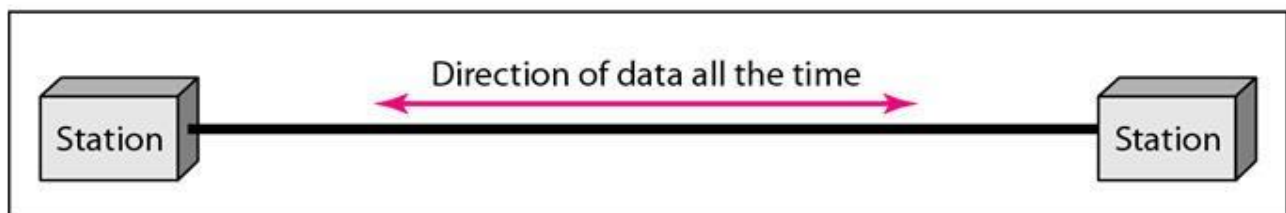
Communication between two devices can be simplex, half-duplex, or full-duplex as shown in Figure



a. Simplex



b. Half-duplex



c. Full-duplex

Simplex:

In simplex mode, the communication is unidirectional, as on a one-way street. Only one of the two devices on a link can transmit; the other can only receive (see Figure a). Keyboards and traditional monitors are examples of simplex devices. The keyboard can only introduce input; the monitor can only accept output. The simplex mode can use the entire capacity of the channel to send data in one direction.

Half-Duplex:

In half-duplex mode, each station can both transmit and receive, but not at the same time. When one device is sending, the other can only receive, and vice versa. The half-duplex mode is like a one-lane road with traffic allowed in both directions.

When cars are traveling in one direction, cars going the other way must wait. In a half-duplex transmission, the entire capacity of a channel is taken over by whichever of the two devices is transmitting at the time. Walkie-talkies and CB (citizens band) radios are both half-duplex systems. The half-duplex mode is used in cases where there is no need for communication in both directions at the same time; the entire capacity of the channel can be utilized for each direction.

Full-Duplex:

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

1.2 NETWORKS

A network is a set of devices (often referred to as *nodes*) connected by communication links. A node can be a computer, printer, or any other device capable of sending and/or receiving data generated by other nodes on the network.

1.2.1 Distributed Processing

Most networks use distributed processing, in which a task is divided among multiple computers. Instead of one single large machine being responsible for all aspects of a process, separate computers (usually a personal computer or workstation) handle a subset.

1.2.2 Network Criteria

A network must be able to meet a certain number of criteria. The most important of these are performance, reliability, and security.

Performance:

Performance can be measured in many ways, including transit time and response time. Transit time is the amount of time required for a message to travel from one device to another. Response time is the elapsed time between an inquiry and a response. The performance of a network depends on a number of factors, including the number of users, the type of transmission medium, the capabilities of the connected hardware, and the efficiency of the software. Performance is often evaluated by two networking metrics: throughput and delay. We often need more throughput and less delay. However, these two criteria are often contradictory. If we try to send more data to the network, we may increase throughput but we increase the delay because of traffic congestion in the network.

Reliability:

In addition to accuracy of delivery, network reliability is measured by the frequency of failure, the time it takes a link to recover from a failure, and the network's robustness in a catastrophe.

Security:

Network security issues include protecting data from unauthorized access, protecting data from damage and development, and implementing policies and procedures for recovery from breaches and data losses.

Scalability increases the system performance by adding more processors

Sharing Resources from one Computer to another Computer over a network

1.2.3 Physical Structures:

Type of Connection

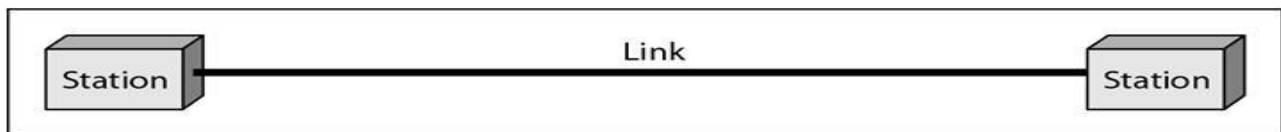
A network is two or more devices connected through links. A link is a communications pathway that transfers data from one device to another. For visualization purposes, it is simplest to imagine any link as a line drawn between two points. For communication to occur, two devices must be connected in some way to the same link at the same time. There are two possible types of connections: point-to-point and multipoint.

Point-to-Point

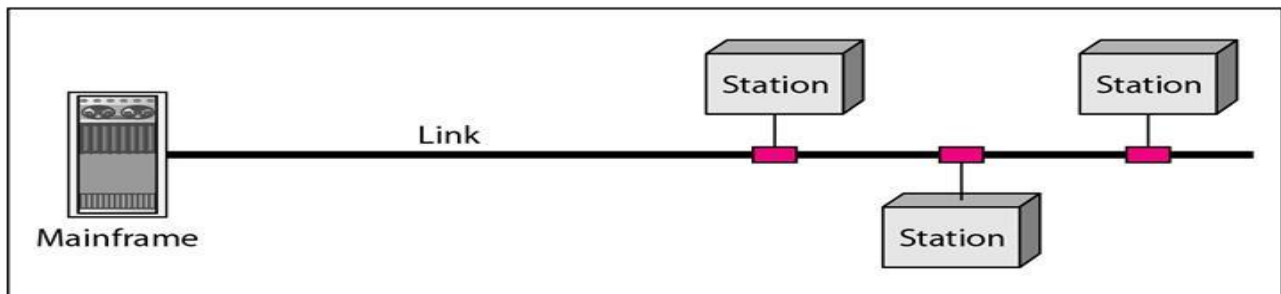
A point-to-point connection provides a dedicated link between two devices. The entire capacity of the link is reserved for transmission between those two devices. Most point-to-point connections use an actual length of wire or cable to connect the two ends, but other options, such as microwave or satellite links, are also possible. When you change television channels by infrared remote control, you are establishing a point-to-point connection between the remote control and the television's control system.

Multipoint

A multipoint (also called multidrop) connection is one in which more than two specific devices share a single link. In a multipoint environment, the capacity of the channel is shared, either spatially or temporally. If several devices can use the link simultaneously, it is a *spatially shared* connection. If users must take turns, it is a *timeshared* connection.



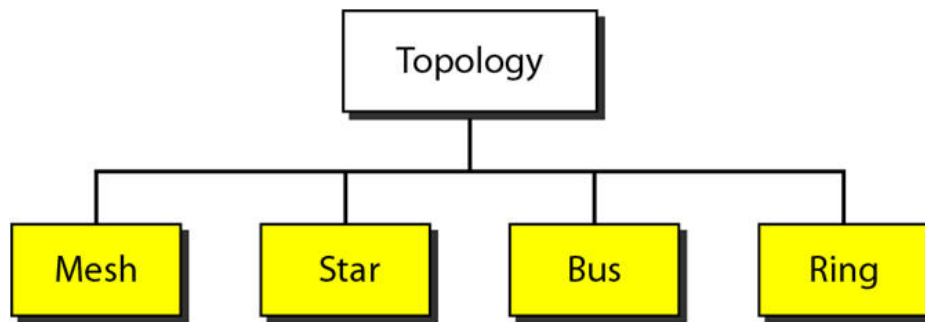
a. Point-to-point



b. Multipoint

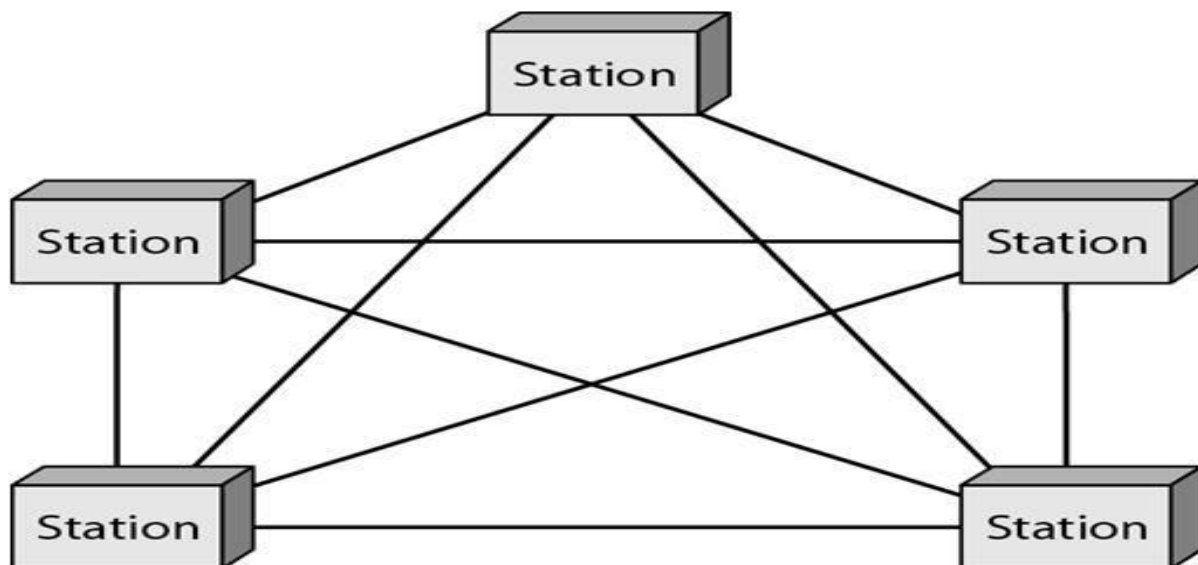
1.2.3.1 Physical Topology

The term *physical topology* refers to the way in which a network is laid out physically. One or more devices connect to a link; two or more links form a topology. The topology of a network is the geometric representation of the relationship of all the links and linking devices (usually called nodes) to one another. There are four basic topologies possible: mesh, star, bus, and ring



Mesh: In a mesh topology, every device has a dedicated point-to-point link to every other device. The term *dedicated* means that the link carries traffic only between the two devices it connects. To find the number of physical links in a fully connected mesh network with n nodes, we first consider that each node must be connected to every other node. Node 1 must be connected to $n - 1$ nodes, node 2 must be connected to $n - 1$ nodes, and finally node n must be connected to $n - 1$ nodes. We need $n(n - 1)$ physical links. However, if each physical link allows communication in both directions (duplex mode), we can divide the number of links by 2. In other words, we can say that in a mesh topology, we need $n(n - 1) / 2$ duplex-mode links.

To accommodate that many links, every device on the network must have $n - 1$ input/output (I/O) ports to be connected to the other $n - 1$ stations.



Advantages:

1. The use of dedicated links guarantees that each connection can carry its own data load, thus eliminating the traffic problems that can occur when links must be shared by multiple devices.
2. A mesh topology is robust. If one link becomes unusable, it does not incapacitate the entire system. Third, there is the advantage of privacy or security. When every message travels along a dedicated line, only the intended recipient sees it. Physical boundaries prevent other users from gaining access to messages. Finally, point-to-point links make fault identification and fault isolation easy. Traffic can be routed to avoid links with suspected problems. This facility enables the network manager to discover the precise location of the fault and aids in finding its cause and solution.

Disadvantages:

1. Disadvantage of a mesh are related to the amount of cabling because every device must be connected to every other device, installation and reconnection are difficult.
2. Second, the sheer bulk of the wiring can be greater than the available space (in walls, ceilings, or floors) can accommodate. Finally, the hardware required to connect each link (I/O ports and cable) can be prohibitively expensive.

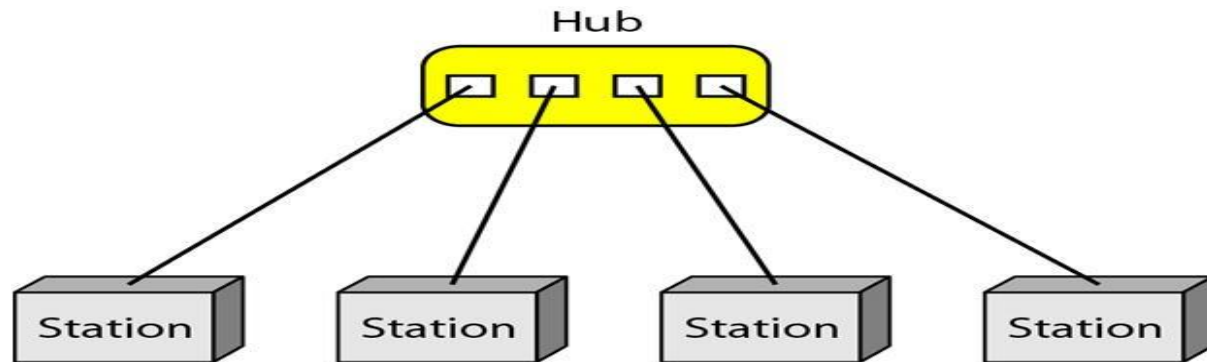
For these reasons a mesh topology is usually implemented in a limited fashion, for example, as a backbone connecting the main computers of a hybrid network that can include several other topologies.

Star Topology:

In a star topology, each device has a dedicated point-to-point link only to a central controller, usually called a hub. The devices are not directly linked to one another. Unlike a mesh topology, a star topology does not allow direct traffic between devices. The controller acts as an exchange: If one device wants to send data to another, it sends the data to the controller, which then relays the data to the other connected device .

A star topology is less expensive than a mesh topology. In a star, each device needs only one link and one I/O port to connect it to any number of others. This factor also makes it easy to install and reconfigure. Far less cabling needs to be housed, and additions, moves, and deletions involve only one connection: between that device and the hub.

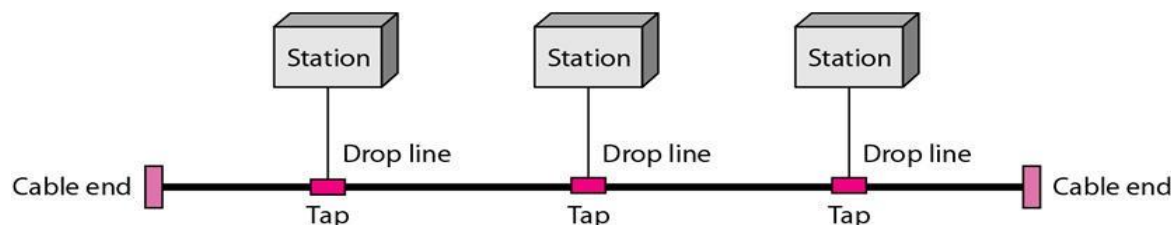
Other **advantages** include robustness. If one link fails, only that link is affected. All other links remain active. This factor also lends itself to easy fault identification and fault isolation. As long as the hub is working, it can be used to monitor link problems and bypass defective links.



One big **disadvantage** of a star topology is the dependency of the whole topology on one single point, the hub. If the hub goes down, the whole system is dead. Although a star requires far less cable than a mesh, each node must be linked to a central hub. For this reason, often more cabling is required in a star than in some other topologies (such as ring or bus).

Bus Topology:

The preceding examples all describe point-to-point connections. A **bus topology**, on the other hand, is multipoint. One long cable acts as a **backbone** to link all the devices in a network



Nodes are connected to the bus cable by drop lines and taps. A drop line is a connection running between the device and the main cable. A tap is a connector that either splices into the main cable or punctures the sheathing of a cable to create a contact with the metallic core. As a signal travels along the backbone, some of its energy is transformed into heat. Therefore, it becomes weaker and weaker as it travels farther and farther. For this reason there is a limit on the number of taps a bus can support and on the distance between those taps.

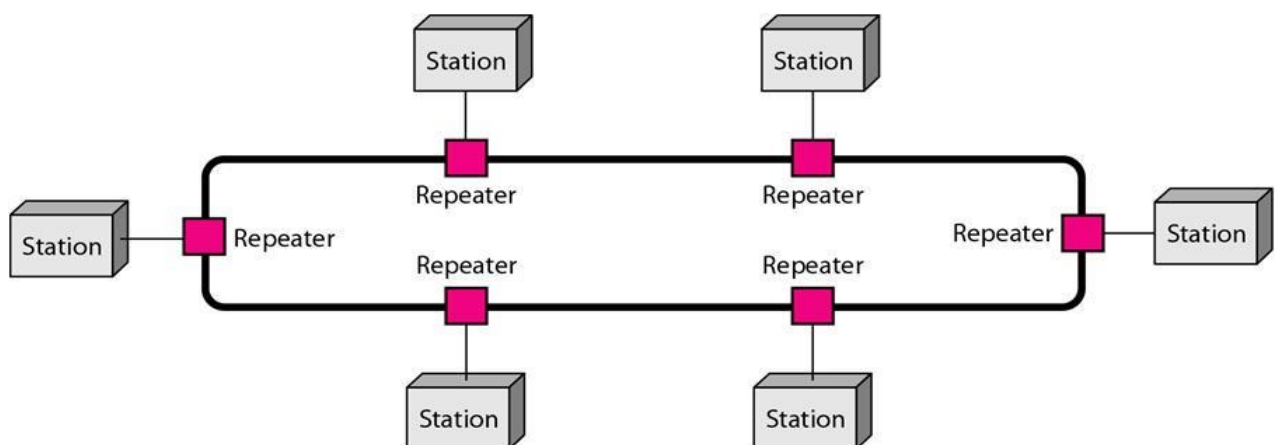
Advantages of a bus topology include ease of installation. Backbone cable can be laid along the most efficient path, then connected to the nodes by drop lines of various lengths. In this way, a bus uses less cabling than mesh or star topologies. In a star, for example, four network devices in the same room require four lengths of cable reaching all the way to the hub. In a bus, this redundancy is eliminated. Only the backbone cable stretches through the entire facility. Each drop line has to reach only as far as the nearest point on the backbone.

Disadvantages include difficult reconnection and fault isolation. A bus is usually designed to be optimally efficient at installation. It can therefore be difficult to add new devices. Signal reflection at the taps can cause degradation in quality. This degradation can be controlled by limiting the number and spacing of devices connected to a given length of cable. Adding new devices may therefore require modification or replacement of the backbone.

In addition, a fault or break in the bus cable stops all transmission, even between devices on the same side of the problem. The damaged area reflects signals back in the direction of origin, creating noise in both directions.

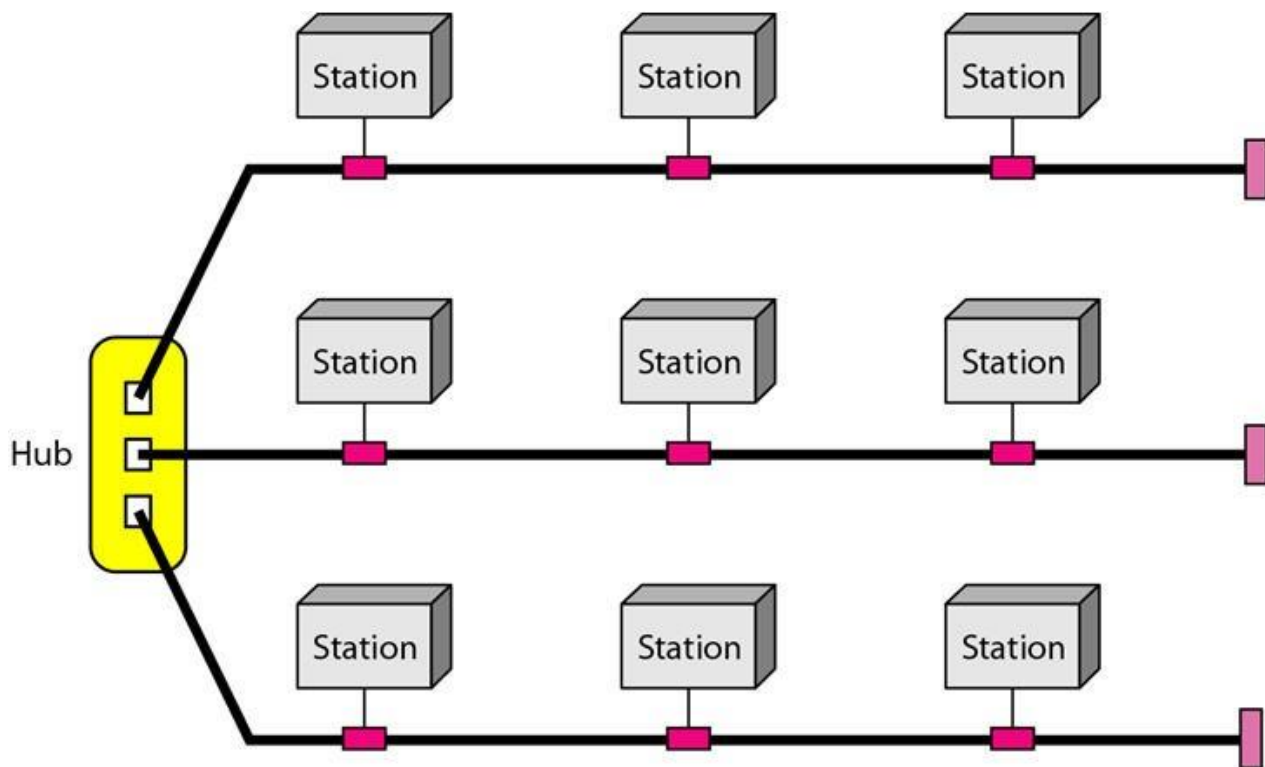
Bus topology was the one of the first topologies used in the design of early local area networks. Ethernet LANs can use a bus topology, but they are less popular.

Ring Topology In a ring topology, each device has a dedicated point-to-point connection with only the two devices on either side of it. A signal is passed along the ring in one direction, from device to device, until it reaches its destination. Each device in the ring incorporates a repeater. When a device receives a signal intended for another device, its repeater regenerates the bits and passes them along



A ring is relatively easy to install and reconfigure. Each device is linked to only its immediate neighbours (either physically or logically). To add or delete a device requires changing only two connections. The only constraints are media and traffic considerations (maximum ring length and number of devices). In addition, fault isolation is simplified. Generally in a ring, a signal is circulating at all times. If one device does not receive a signal within a specified period, it can issue an alarm. The alarm alerts the network operator to the problem and its location.

However, unidirectional traffic can be a disadvantage. In a simple ring, a break in the ring (such as a disabled station) can disable the entire network. This weakness can be solved by using a dual ring or a switch capable of closing off the break. Ring topology was prevalent when IBM introduced its local-area network Token Ring. Today, the need for higher-speed LANs has made this topology less popular. Hybrid Topology A network can be hybrid. For example, we can have a main star topology with each branch connecting several stations in a bus topology as shown in Figure



1.2.4 Categories of Networks

Local Area Networks:

Local area networks, generally called LANs, are privately-owned networks within a single building or campus of up to a few kilometres in size. They are widely used to connect personal computers and workstations in company offices and factories to share resources (e.g., printers) and exchange information. LANs are distinguished from other kinds of networks by three characteristics:

- (1) Their size,
- (2) Their transmission technology, and
- (3) Their topology.

LANs are restricted in size, which means that the worst-case transmission time is bounded and known in advance. Knowing this bound makes it possible to use certain kinds of designs that would not otherwise be possible. It also simplifies network management. LANs may use a transmission technology consisting of a cable to which all the machines are attached, like the telephone company party lines once used in rural areas. Traditional LANs run at speeds of 10 Mbps to 100 Mbps, have low delay (microseconds or nanoseconds), and make very few errors. Newer LANs operate at up to 10 Gbps. Various topologies are possible for broadcast LANs. Figure 1 shows two of them. In a bus (i.e., a linear cable) network, at any instant at most one machine is the master and is allowed to transmit. All other machines are required to refrain from sending. An arbitration mechanism is needed to resolve conflicts when two or more machines want to transmit simultaneously. The arbitration mechanism may be centralized or distributed. IEEE 802.3, popularly called Ethernet, for example, is a bus-based broadcast network with decentralized control, usually operating at 10 Mbps to 10 Gbps. Computers on an Ethernet can transmit whenever they want to; if two or more packets collide, each computer just waits a random time and tries again later.

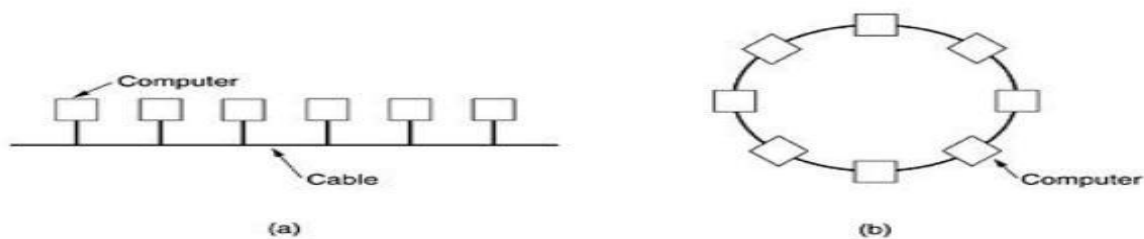


Fig.1: Two broadcast networks . (a) Bus. (b) Ring.

A second type of broadcast system is the ring. In a ring, each bit propagates around on its own, not waiting for the rest of the packet to which it belongs. Typically, each bit circumnavigates the entire ring in the time it takes to transmit a few bits, often before the complete packet has even been transmitted. As with all other broadcast systems, some rule is needed for arbitrating simultaneous accesses to the ring. Various methods, such as having the machines take turns, are in use. IEEE 802.5 (the IBM token ring), is a ring-based LAN operating at 4 and 16 Mbps. FDDI is another example of a ring network.

Metropolitan Area Network (MAN):

Metropolitan Area Network:

A metropolitan area network, or MAN, covers a city. The best-known example of a MAN is the cable television network available in many cities. This system grew from earlier community antenna systems used in areas with poor over-the-air television reception. In these early systems, a large antenna was placed on top of a nearby hill and signal was then piped to the subscribers' houses. At first, these were locally-designed, ad hoc systems. Then companies began jumping into the business, getting contracts from city governments to wire up an entire city. The next step was television programming and even entire channels designed for cable only. Often these channels were highly specialized, such as all news, all sports, all cooking, all gardening, and so on. But from their inception until the late 1990s, they were intended for television reception only. To a first approximation, a MAN might look something like the system shown in Fig. In this figure both television signals and Internet are fed into the centralized head end for subsequent distribution to people's homes. Cable television is not the only MAN. Recent developments in high-speed wireless Internet access resulted in another MAN, which has been standardized as IEEE 802.16.

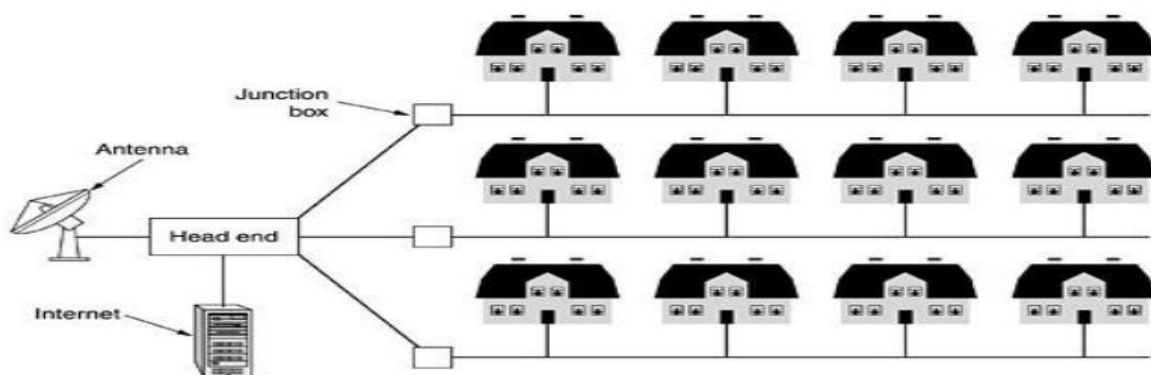


Fig.2: Metropolitan area network based on cable TV.

A MAN is implemented by a standard called DQDB (Distributed Queue Dual Bus) or IEEE 802.16. DQDB has two unidirectional buses (or cables) to which all the computers are attached.

Wide Area Network (WAN).

Wide Area Network:

A wide area network, or WAN, spans a large geographical area, often a country or continent. It contains a collection of machines intended for running user (i.e., application) programs. These machines are called as hosts. The hosts are connected by a communication subnet, or just subnet for short. The hosts are owned by the customers (e.g., people's personal computers), whereas the communication subnet is typically owned and operated by a telephone company or Internet service provider. The job of the subnet is to carry messages from host to host, just as the telephone system carries words from speaker to listener.

Separation of the pure communication aspects of the network (the subnet) from the application aspects (the hosts), greatly simplifies the complete network design. In most wide area networks, the subnet consists of two distinct components: transmission lines and switching elements. Transmission lines move bits between machines. They can be made of copper wire, optical fiber, or even radio links. In most WANs, the network contains numerous transmission lines, each one connecting a pair of routers. If two routers that do not share a transmission line wish to communicate, they must do this indirectly, via other routers. When a packet is sent from one router to another via one or more intermediate routers, the packet is received at each intermediate router in its entirety, stored there until the required output line is free, and then forwarded. A subnet organized according to this principle is called a store-and-forward or packet-switched subnet. Nearly all wide area networks (except those using satellites) have store-and-forward subnets. When the packets are small and all the same size, they are often called cells.

The principle of a packet-switched WAN is so important. Generally, when a process on some host has a message to be sent to a process on some other host, the sending host first cuts the message into packets, each one bearing its number in the sequence. These packets are then injected into the network one at a time in quick succession. The packets are transported individually over the network and deposited at the receiving host, where they are reassembled into the original message and delivered to the receiving process. A stream of packets resulting from some initial message is illustrated in Fig.

In this figure, all the packets follow the route ACE, rather than ABDE or ACDE. In some networks all packets from a given message must follow the same route; in others each packet is routed separately. Of course, if ACE is the best route, all packets may be sent along it, even if each packet is individually routed.

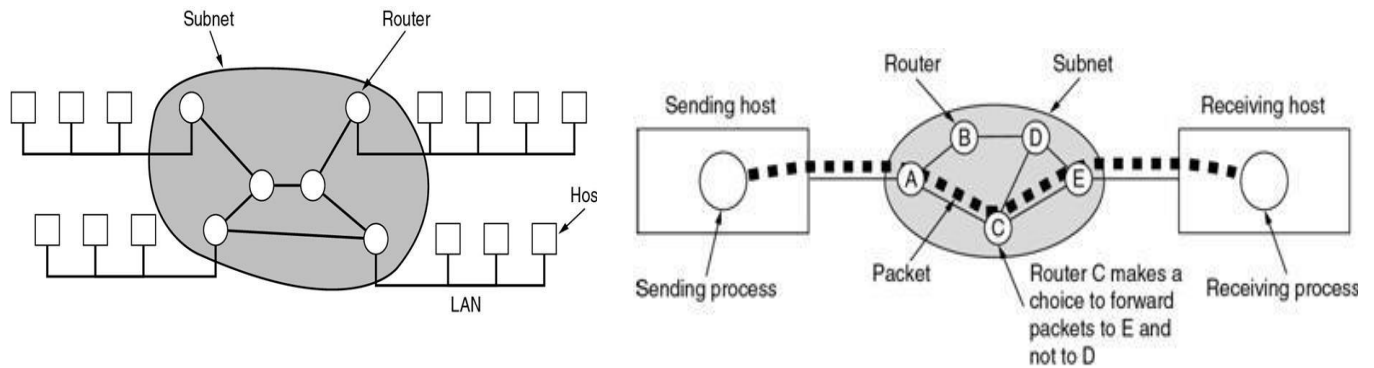
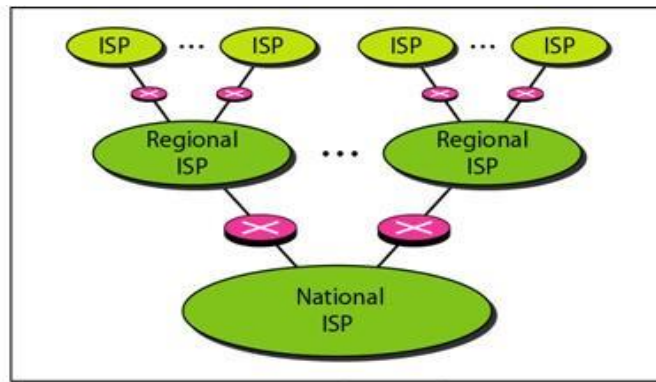


Fig.3.1: A stream of packets from sender to receiver.

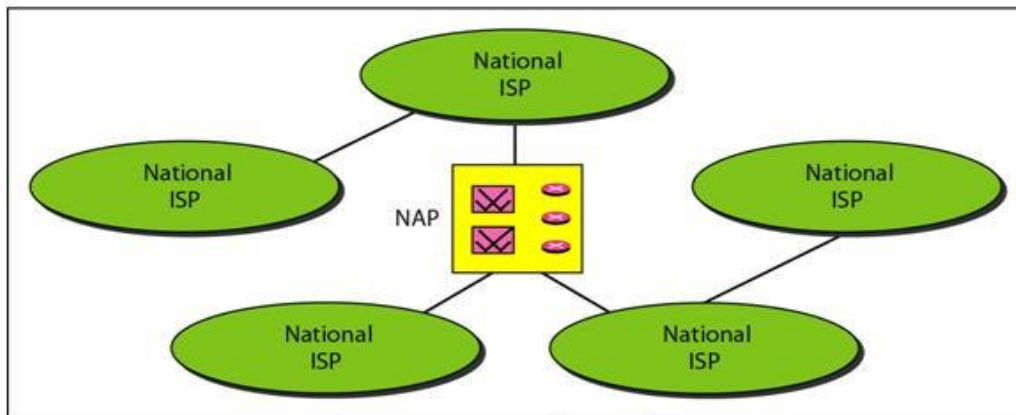
Not all WANs are packet switched. A second possibility for a WAN is a satellite system. Each router has an antenna through which it can send and receive. All routers can hear the output from the satellite, and in some cases they can also hear the upward transmissions of their fellow routers to the satellite as well. Sometimes the routers are connected to a substantial point-to-point subnet, with only some of them having a satellite antenna. Satellite networks are inherently broadcast and are most useful when the broadcast property is important.

The Internet Today

The Internet has come a long way since the 1960s. The Internet today is not a simple hierarchical structure. It is made up of many wide- and local-area networks joined by connecting devices and switching stations. It is difficult to give an accurate representation of the Internet because it is continually changing-new networks are being added, existing networks are adding addresses, and networks of defunct companies are being removed. Today most end users who want Internet connection use the services of Internet service providers (ISPs). There are international service providers, national service providers, regional service providers, and local service providers. The Internet today is run by private companies, not the government. Figure 1.13 shows a conceptual (not geographic) view of the Internet.



a. Structure of a national ISP



b. Interconnection of national ISPs

International Internet Service Providers:

At the top of the hierarchy are the international service providers that connect nations together.

National Internet Service Providers:

The national Internet service providers are backbone networks created and maintained by specialized companies. There are many national ISPs operating in North America; some of the most well known are SprintLink, PSINet, UUNet Technology, AGIS, and internet Mel. To provide connectivity between the end users, these backbone networks are connected by complex switching stations (normally run by a third party) called network access points (NAPs). Some national ISP networks are also connected to one another by private switching stations called *peering points*. These normally operate at a high data rate (up to 600 Mbps).

Regional Internet Service Providers:

Regional internet service providers or regional ISPs are smaller ISPs that are connected to one or more national ISPs. They are at the third level of the hierarchy with a smaller data rate. **Local Internet Service**

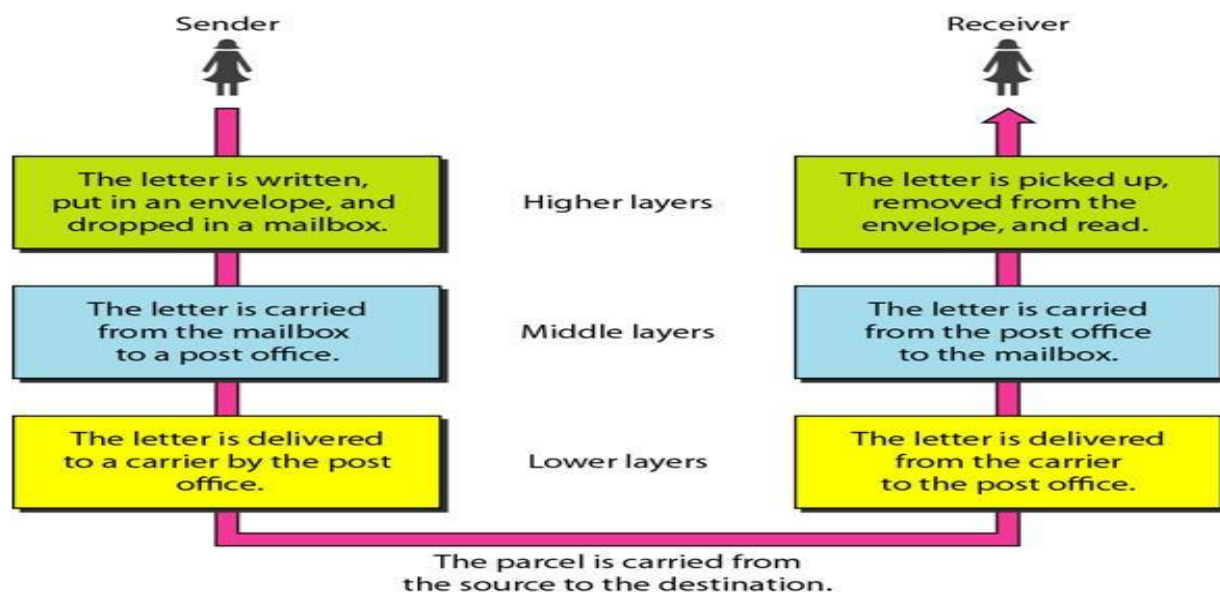
Providers:

Local Internet service providers provide direct service to the end users. The local ISPs can be connected to regional ISPs or directly to national ISPs. Most end users are connected to the local ISPs. Note that in this sense, a local ISP can be a company that just provides Internet services, a corporation with a network that

supplies services to its own employees, or a nonprofit organization, such as a college or a university, that runs its own network. Each of these local ISPs can be connected to a regional or national service provider.

1.3 LAYERED TASKS

We use the concept of layers in our daily life. As an example, let us consider two friends who communicate through postal mail. The process of sending a letter to a friend would be complex if there were no services available from the post office. Below Figure shows the steps in this task.



Sender, Receiver, and Carrier

In Figure we have a sender, a receiver, and a carrier that transports the letter. There is a hierarchy of tasks.

At the Sender Site

Let us first describe, in order, the activities that take place at the sender site.

- o Higher layer. The sender writes the letter, inserts the letter in an envelope, writes the sender and receiver addresses, and drops the letter in a mailbox.

- o Middle layer. The letter is picked up by a letter carrier and delivered to the post office.

- o Lower layer. The letter is sorted at the post office; a carrier transports the letter.

On the Way: The letter is then on its way to the recipient. On the way to the recipient's local post office, the letter may actually go through a central office. In addition, it may be transported by truck, train, airplane, boat, or a combination of these.

At the Receiver Site

- o Lower layer. The carrier transports the letter to the post office.
- o Middle layer. The letter is sorted and delivered to the recipient's mailbox.
- o Higher layer. The receiver picks up the letter, opens the envelope, and reads it.

1.4.1 Connection Oriented and Connectionless Services

These are the two services given by the layers to layers above them.

These services are :

- Connection Oriented Service
- Connectionless Services

Connection Oriented Services

The communication proceeds through three well - defined phases:

- Connection establishment
- Data transfer
- Connection is release

In connection oriented service to establish a connection before starting the communication. When connection is established send the message or the information and then release the connection. Connection oriented service is more reliable than connectionless service. Resend the message in connection oriented service if there is an error at the receivers end. Example of connection oriented is TCP (Transmission Control Protocol) protocol.

Connection less Services

it carries the full address where the message (letter) is to be carried. Each message is routed independently from source to destination. The order of message sent can be different from the order received. It is similar to the postal services

Criteria	Connection-Oriented	Connection-Less
Definition	It is the communication system in which virtual connection is established between sender and receiver before the communication begins	It is the Communication System in which there is no need to establish virtual connection between sender and receiver.
Connection	prior connection needs to be established	No prior connection is established
Resource allocation	need to be allocated	No prior allocation of resource is required.
Data Acknowledgement	Receiver can acknowledge the data send by the sender and can re request the data if any packet fails or gets	No data acknowledge is used, sender cannot be sure about the accurate delivery of the message.

	damaged	
Reliability	it ensures reliable Transfer of data	Reliability is not Guaranteed as it is a best effort service
Transfer Mode	it can implemented either circuit switching or Virtual Circuit.	it is implemented using Packet switching.
Packet Route	All the frames are sent through same route or path.	Packets follow different path to reach destination and may reach in any order
Retransmission	it is possible to retransmit the lost data bits	it is not possible to retransmit the lost data bits .
Delay	There is more delay in transfer of information ,but once connection established faster delivery	There is no delay due to absence of connection establishment.
Congestion	Congestion is not at all possible	Congestion can occur likely
Connection Termination	Connection needs to be terminated after completion of communication	No Need of Connection termination
Example	Telephone Call	Postal System
	Example of connection oriented is TCP (Transmission Control Protocol)	Example of Connectionless service is UDP (User Datagram Protocol)
	x.25,frame relay	Datagram packet switching

1.4 The OSI Reference Model:

The OSI model (minus the physical medium) is shown in Fig. This model is based on a proposal developed by the International Standards Organization (ISO) as a first step toward international standardization of the protocols used in the various layers (Day and Zimmermann, 1983). It was revised in 1995(Day, 1995). The model is called the ISO-OSI (Open Systems Interconnection) Reference Model because it deals with connecting open systems—that is, systems that are open for communication with other systems.

The OSI model has seven layers. The principles that were applied to arrive at the seven layers can be briefly summarized as follows:

1. A layer should be created where a different abstraction is needed.
2. Each layer should perform a well-defined function.
3. The function of each layer should be chosen with an eye toward defining internationally standardized protocols.
4. The layer boundaries should be chosen to minimize the information flow across the interfaces.
5. The number of layers should be large enough that distinct functions need not be thrown together in the same layer out of necessity and small enough that the architecture does not become unwieldy.

Organization of the layers

The 7 layers can be grouped into 3 subgroups

Network Support Layers

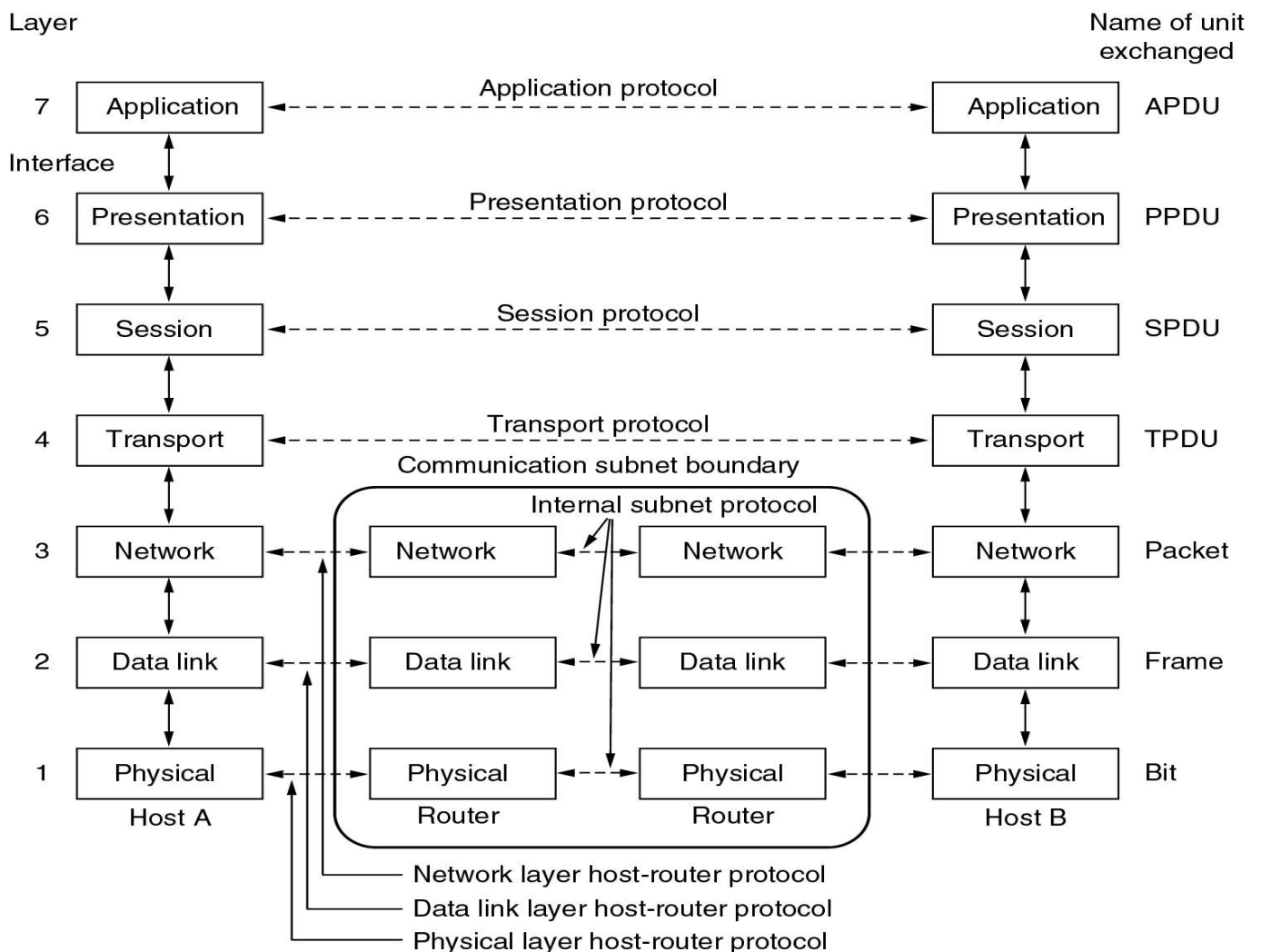
Layers 1,2,3 - Physical, Data link and Network are the network support layers. They deal with the physical aspects of moving data from one device to another such as electrical specifications, physical addressing, transport timing and reliability.

Transport Layer

Layer 4, transport layer, ensures end-to-end reliable data transmission on a single link.

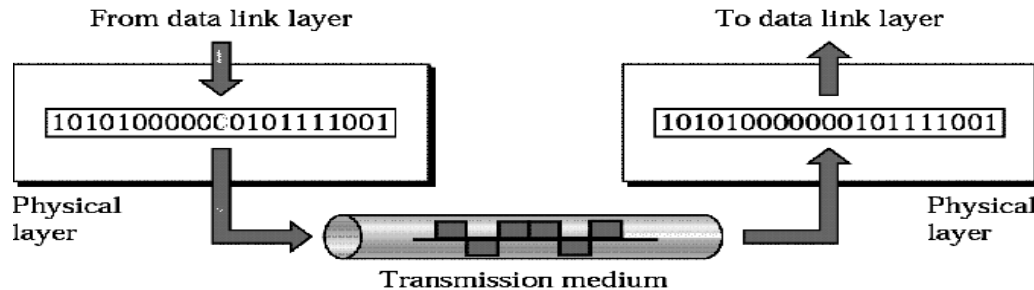
User Support Layers

Layers 5,6,7 – Session, presentation and application are the user support layers. They allow interoperability among unrelated software systems



Physical Layer

The physical layer coordinates the functions required to transmit a bit stream over a physical medium.



The physical layer is 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 medium.

Representation of bits - To transmit the stream of bits, it must be encoded to signals. The physical layer defines the type of encoding.

Data Rate or Transmission rate - The number of bits sent each second – is also defined by the physical layer.

Synchronization of bits - The sender and receiver must be synchronized at the bit level. Their clocks must be synchronized.

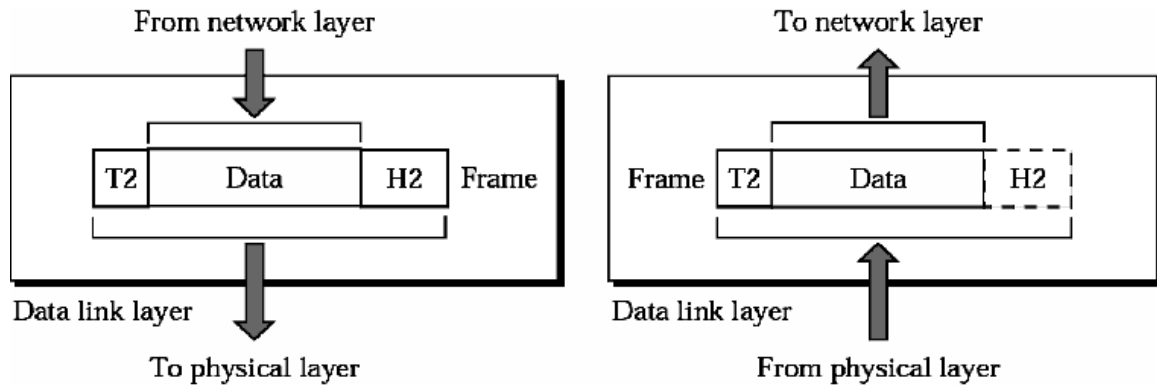
Line Configuration - In a point-to-point configuration, two devices are connected together through a dedicated link. In a multipoint configuration, a link is shared between several devices.

Physical Topology - The physical topology defines how devices are connected to make a network. Devices can be connected using a mesh, bus, star or ring topology.

Transmission Mode - The physical layer also defines the direction of transmission between two devices: simplex, half-duplex or full-duplex.

Data Link Layer

It is responsible for transmitting frames from one node to next node.



The other responsibilities of this layer are

Framing - Divides the stream of bits received into data units called frames.

Physical addressing – If frames are to be distributed to different systems on the n/w , data link layer adds a header to the frame to define the sender and receiver.

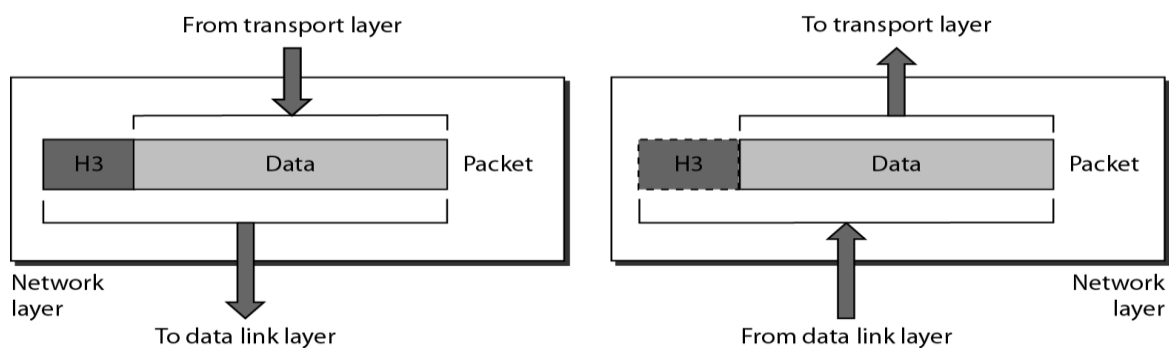
Flow control- If the rate at which the data are absorbed by the receiver is less than the rate produced in the sender ,the Data link layer imposes a flow ctrl mechanism.

Error control- Used for detecting and retransmitting damaged or lost frames and to prevent duplication of frames. This is achieved through a trailer added at the end of the frame.

Access control -Used to determine which device has control over the link at any given time.

The Network Layer:

The network layer is responsible for the source-to-destination delivery of a packet, possibly across multiple networks (links).



It is mainly required, when it is necessary to send information from one network to another.

The other responsibilities of this layer are

Logical addressing - If a packet passes the n/w boundary, we need another addressing system for source and destination called logical address.

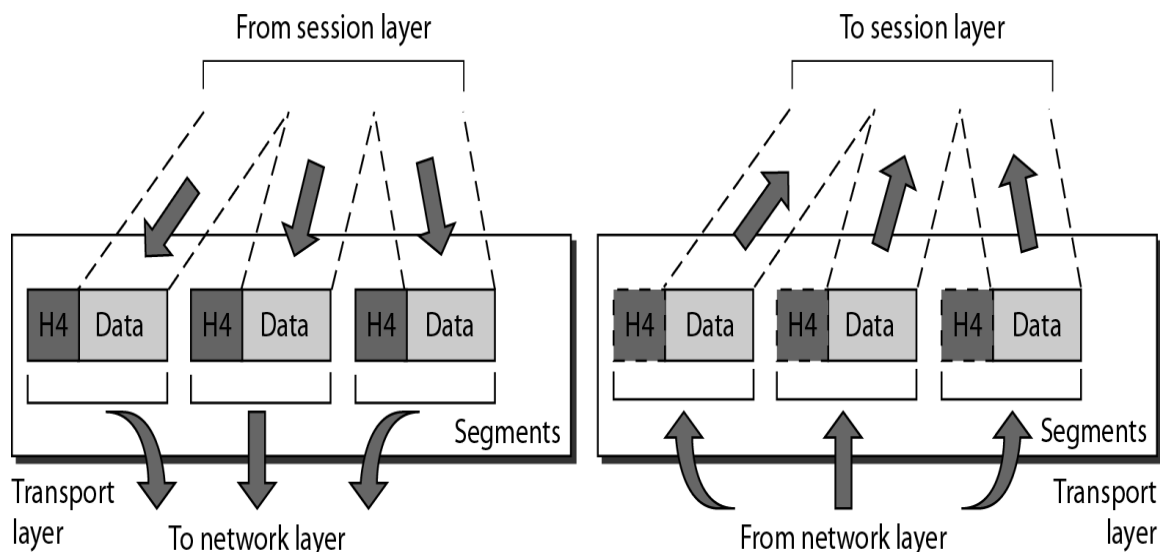
Routing – The devices which connects various networks called routers are responsible for delivering packets to final destination.

When independent networks or links are connected to create **internetworks** or a large network, the connecting devices (called *routers* or *switches*) route or switch the packets to their final destination.

Transport Layer

It is responsible for **Process to Process** delivery.

It also ensures whether the message arrives in order or not.



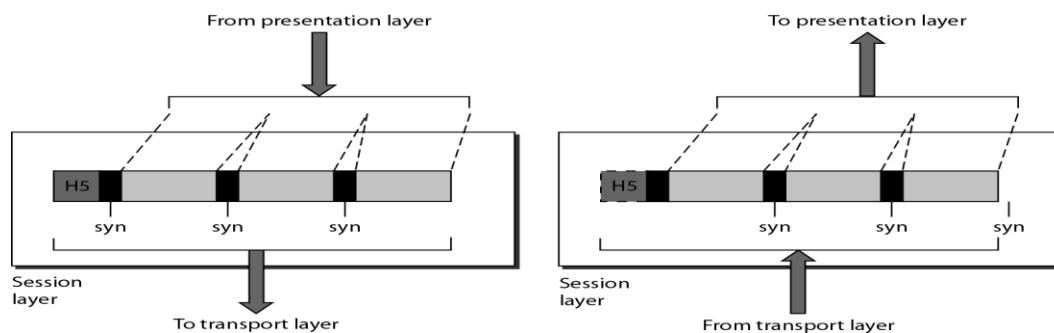
Port addressing - The header in this must therefore include a address called port address. This layer gets the entire message to the correct process on that computer.

Segmentation and reassembly - The message is divided into segments and each segment is assigned a sequence number. These numbers are arranged correctly on the arrival side by this layer.

Connection control - This can either be **connectionless** or **connection-oriented**. The connectionless treats each segment as a individual packet and delivers to the destination. The connection-oriented makes connection on the destination side before the delivery. After the delivery the termination will be terminated.

Flow and error control - Similar to data link layer, but process to process take place.

Session Layer



This layer establishes, manages and terminates connections between applications.

Specific responsibilities of the session layer include the following:

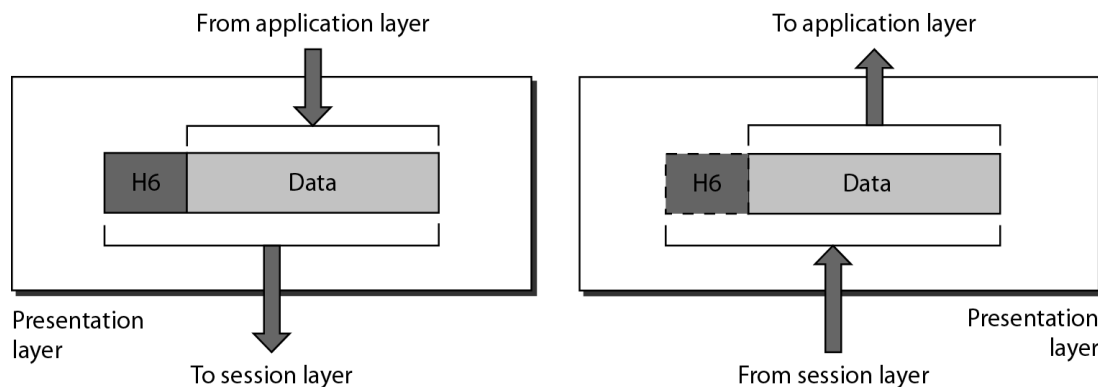
Dialog control: Keeping track of whose turn it is to transmit. The session layer allows two systems to enter into a dialog. It allows the communication between two processes to take place in either half duplex (one way at a time) or full-duplex (two ways at a time) mode.

Synchronization: The session layer allows a process to add checkpoints, or synchronization Points, to a stream of data. Check pointing long transmissions to allow them to pick up from where they left off in the event of a crash and subsequent recovery.

Token management: Preventing two parties from attempting the same critical operation simultaneously.

Presentation Layer

It is concerned with the syntax and semantics of information exchanged between two systems.



The other responsibilities of this layer are

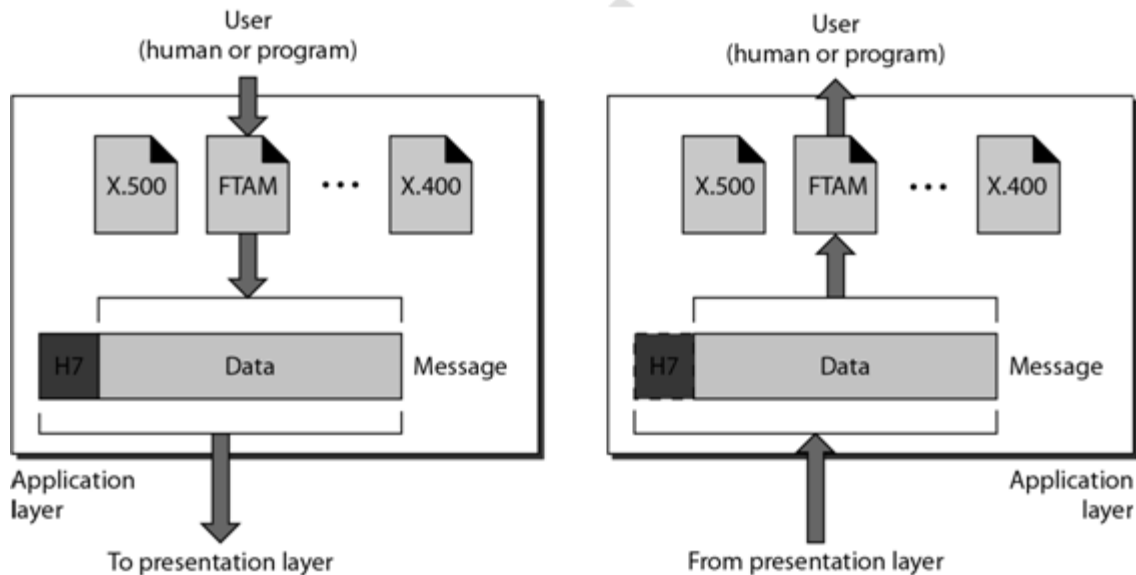
Translation – Different computers use different encoding system, this layer is responsible for interoperability between these different encoding methods. It will change the message into some common format.

Encryption and decryption–It means that sender transforms the original information to another form and sends the resulting message over the n/w. and vice versa.

Compression and expansion–Compression reduces the number of bits contained in the information particularly in text, audio and video.

Application Layer

This layer enables the user to access the n/w. This allows the user to log on to remote user. Network virtual terminal.



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

File transfer, access, and management.

This application allows a user to access files in a remote host (to make changes or read data), to retrieve files from a remote computer for use in the local computer, and to manage or control files in a remote computer locally.

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

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

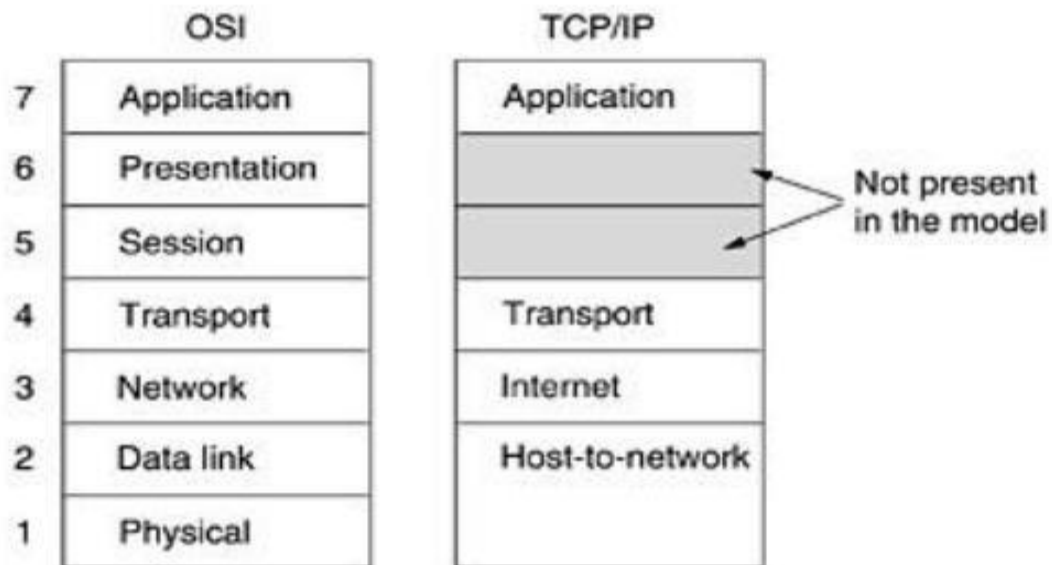
1.7.The TCP/IP Reference Model:

The TCP/IP reference model was developed prior to OSI model. The major design goals of this model were,

1. To connect multiple networks together so that they appear as a single network.
2. To survive after partial subnet hardware failures.
3. To provide a flexible architecture.

Unlike OSI reference model, TCP/IP reference model has only 4 layers. They are,

1. The Link Layer or Host-to-Network Layer
2. The Internet Layer
3. The Transport Layer
4. The Application Layer



TCP/IP PROTOCOL SUITE

TCP/IP is a hierarchical protocol made up of interactive modules, each of which Provides a specific functionality.

Link Layer or Host-to-Network Layer

TCP/IP does not define any specific protocol. It supports all the standard and proprietary protocols. A network in a TCP/IP internetwork can be a local-area network or a wide-area network. The lowest layer in the model, the **link layer** describes what links such as serial lines and classic Ethernet must do to meet the needs of this connectionless internet layer. It is not really a layer at all, in the normal sense of the term, but rather an interface between Hosts and transmission links.

The Internet Layer

At the network layer (or, more accurately, the internetwork layer), TCP/IP supports the Internetworking Protocol. IP, four supporting protocols: ARP, RARP, ICMP, and IGMP. Its job is to permit hosts to inject packets into any network and have them travel independently to the destination (potentially on a different network). They may even arrive in a completely different order than they were sent, in which case it is the job of higher layers to rearrange them, if in-order delivery is desired. The internet layer defines an official packet format and protocol called **IP (Internet Protocol)**

Internetworking Protocol (IP)

The Internetworking Protocol (IP) is the transmission mechanism used by the TCP/IP protocols. It is an unreliable and connectionless protocol—a best-effort delivery service. The term best effort means that IP provides no error checking or tracking. IP assumes the unreliability of the underlying layers and does its best to get a transmission through to its destination, but with no guarantees. IP transports data in packets called data grams, each of which is transported separately. Data grams 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 data grams Once they arrive at their destination.

Address Resolution Protocol

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

Reverse Address Resolution Protocol

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

Internet Control Message Protocol

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

Internet Group Message Protocol

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

The Transport Layer

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

User Datagram Protocol

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

Transmission Control Protocol

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

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

transmission based on sequence numbers.

Stream Control Transmission Protocol

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

The Application Layer

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

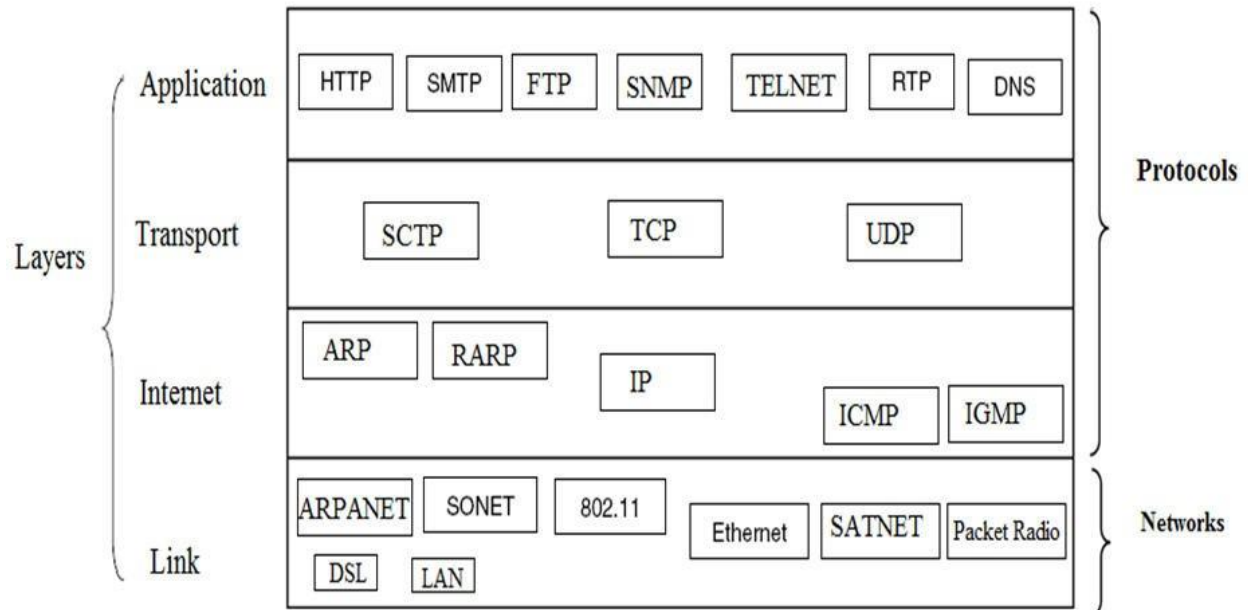


Fig.2: Protocols and networks in the TCP/IP model initially.

The TCP/IP model does not have session or presentation layers. On top of the transport layer is the application layer. It contains all the higher-level protocols. The early ones included virtual terminal (TELNET), file transfer (FTP), and electronic mail (SMTP), as shown in Fig.6.2. The virtual terminal protocol allows a user on one machine to log onto a distant machine and work there. The file transfer protocol provides a way to move data efficiently from one machine to another. Electronic mail was originally just a kind of file transfer, but later a specialized protocol (SMTP) was developed for it. Many other protocols have been added to these over the years: the Domain Name System (DNS) for mapping host names onto their network addresses, NNTP, the protocol for moving USENET news articles around, and HTTP, the protocol for fetching pages on the World Wide Web, and many others.

1.8 Comparison of the OSI and TCP/IP Reference Models:

Following are some major differences between OSI Reference Model and TCP/IP Reference Model, with diagrammatic comparison below.

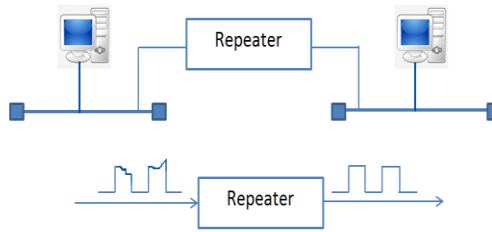
OSI(Open System Interconnection)	TCP/IP(Transmission Control Protocol / Internet Protocol)
1. OSI is a generic, protocol independent standard, acting as a communication gateway between the network and end user.	1. TCP/IP model is based on standard protocols around which the Internet has developed. It is a communication protocol, which allows connection of hosts over a network.
2. OSI model defines services, interfaces and protocols very clearly and makes clear distinction between them. It is protocol independent.	2. In TCP/IP, services, interfaces and protocols are not clearly separated. It is also protocol dependent.
3. It has 7 layers	3. It has 4 layers
4. Network layer of OSI model provides both connection oriented and connectionless service.	4. The Network layer in TCP/IP model provides connectionless service.
5. In OSI model the transport layer guarantees the delivery of packets.(connection oriented communication supports in transport layer)	5. In TCP/IP model the transport layer does not guarantees delivery of packets. Still the TCP/IP model is more reliable. (connection oriented and connection less communication supports in transport layer)
6. OSI model has a separate Presentation layer and Session layer.	6. TCP/IP does not have a separate Presentation layer or Session layer.
7. OSI is a reference model around which the networks are built. Generally it is used as a guidance tool.	7. TCP/IP model is, in a way implementation of the OSI model.
8. OSI model has a problem of fitting the protocols into the model.	8. TCP/IP model does not fit any protocol
9. Protocols are hidden in OSI model and are easily replaced as the technology changes.	9. In TCP/IP replacing protocol is not easy.
10.OSI Model follows a bottom-up approach.	10.TCP/IP follows top to bottom approach
11. Follows vertical approach.	11. Follows horizontal approach.
12.OSI developed model then protocol.	12.TCP/IP developed protocols then model.
13.OSI is protocol independent.	13.TCP/IP is protocol dependent.

Network Devices (Hub, Repeater, Bridge, Switch, Router, Gateways and Brouter)

Network devices, also known as networking hardware, are physical devices that allow hardware on a computer network to communicate and interact with one another. For example Repeater, Hub, Bridge, Switch, Routers, Gateway, Brouter, and NIC, etc.

Repeater :

- The data signals generally becomes too weak or corrupted if they tend travel a long distance.
- Repeater regenerates the signal over the same network.
- A repeater operates at the physical layer or layer 1.
- They do not amplify the signal.
- It is a 2-port device.



Hub :

- A Hub operates at the physical layer or layer 1.
- A hub is a basically multi-port repeater.
- It also known as network Hub
- Hub works at the physical layer of the OSI Model.
- Used to set up LAN
- Hub has no memory
- The connector in star topology which connects different stations.
- Hubs cannot filter data, so data packets are sent to all connected devices.
- When a packet arrives at one port ,it is copied to other ports so that all segments of the LAN can see all packets.



Hub Pros and Cons

PROS:

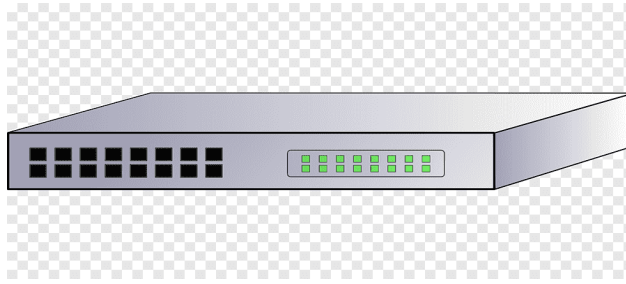
- cheaper than switches
- Works good for smaller network.

CONS:

- Issues with broadcast.
- No memory
- Normally runs in half duplex Mode

Switch

- Switch is layer 2 Device or switch operates in data link layer.
- A switch is a hardware device that connects multiple devices on a computer network to establish a LAN
- The Switch contains the updated table that decides where the data is transmitted or not.
- Switch delivers the message to the correct destination based on the physical address present in the incoming message.
- Switch has memory.
- Stores and maintain the MAC addresses table.



HUB vs Switch

HUB	SWITCH
Layer 1 device	Layer 2 device
Works at Physical layer	Works at data link layer
Has no memory	Has memory and stores MAC address Table.
Not an intelligent Device.	Intelligent Device.
Floods the network due to broadcasting.	Can do unicasting, Multicasting, Broadcasting.
Security risks are high	Security risks are low
Less efficient	More efficient
Half duplex	Full duplex

Bridge –

- A bridge operates at the data link layer.
- A bridge is a repeater, with add on the functionality of filtering content by reading the MAC addresses of
- the source and destination.
- It is also used for interconnecting two LANs working on the same protocol.
- It has a single input and single output port, thus making it a 2 port device.

Types of Bridges

Transparent Bridges:-

- These are the bridge in which the stations are completely unaware of the bridge's existence i.e. whether or not a bridge is added or deleted from the network, reconfiguration of the stations is unnecessary.
- These bridges make use of two processes i.e. bridge forwarding and bridge learning.

Source Routing Bridges:-

- In these bridges, routing operation is performed by the source station and the frame specifies which route to follow.
- The host can discover the frame by sending a special frame called the discovery frame, which spreads through the entire network using all possible paths to the destination.

Router

- A router is a hardware device which is used to connect a LAN with an internet connection.
- Connects two or more different LANs.
- A router works in a Layer 3 (Network layer) of the OSI Reference model.
- Store s Routing Table.
- It is used to receive, analyze and forward the incoming packets to another network.
- A router forwards the packet based on the information available in the routing table.
- It determines the best path from the available paths for the transmission of the packet.
