UNIT-1
Computer Networks and the Internet

1. What Is the Internet?

The Internet is a computer network that interconnects hundreds of millions of computing devices throughout the world.

A Nuts-and-Bolts Description

Computer networking is a collection of computers and other devices that can exchange data and share resources with each other.

The devices are called hosts or end systems, devices were primarily traditional desktop PCs, workstations, servers, smart phones, laptops etc.,

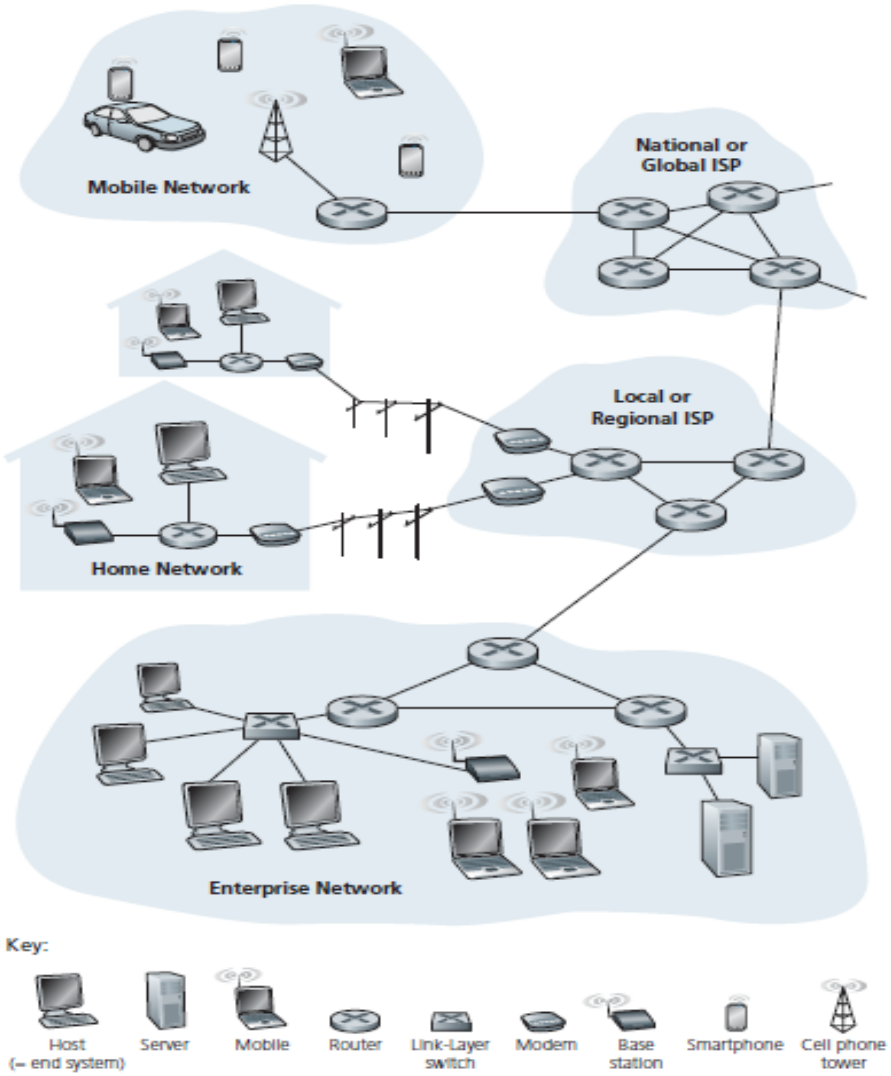

Fig: Some pieces of the Internet

- End systems are connected together by a network of **communication links** and **packet switches**.
- there are many types of communication links, which are made up of different types of physical media, including coaxial cable, copper wire, optical fiber, and radio spectrum.
- Different links can transmit data at different rates, with the **transmission rate** of a

link measured in bits/second.

- When one end system has data to send to another end system, the sending end system segments the data and adds header bytes to each segment. The resulting packages of information, known as **packets.**
- A packet switch takes a packet arriving on one of its incoming communication links and forwards that packet on one of its outgoing communication links. The two most prominent types in today's Internet are **routers** and **link-layer switches**. Both types of switches forward packets toward their ultimate destinations.
- The sequence of communication links and packet switches traversed by a packet from the sending end system to the receiving end system is known as a **route** or **path.**
- End systems access the Internet through **Internet Service Providers (ISPs),**including residential ISPs such as local cable or telephone companies; corporate ISPs; university ISPs; and ISPs that provide WiFi access in airports, hotels, coffee shops, and other public places.
- End systems, packet switches, and other pieces of the Internet run **protocols** that control the sending and receiving of information within the Internet. The **Transmission Control Protocol (TCP)** and the **Internet Protocol (IP)** are two of the most important protocols in the Internet.
- The IP protocol specifies the format of the packets that are sent and received among routers and end systems. The Internet's principal protocols are collectively known as **TCP/IP**.

A Services Description

- As *an infrastructure that provides services to applications*. These applications include electronic mail, Web surfing, social networks, instant messaging, video streaming, distributed games, peer-to-peer (P2P) file sharing, television over the Internet, remote login, and much, much more.
- The applications are said to be **distributed applications**, since they involve multiple end systems that exchange data with each other. Importantly, Internet applications run on end systems—they do not run in the packet switches in the network core.
- End systems attached to the Internet provide an **Application Programming Interface (API)** that specifies how a program running on one end system asks the Internet infrastructure to deliver data to a specific destination program running on another end system.
- This Internet API is a set of rules that the sending program must follow so that the Internet can deliver the data to the destination program.

What Is a Protocol?

A **protocol** defines the format and the order of messages exchanged between two or more communicating entities, as well as the actions taken on the transmission and/or receipt of a message or other event.

A Human Analogy and Network Protocols

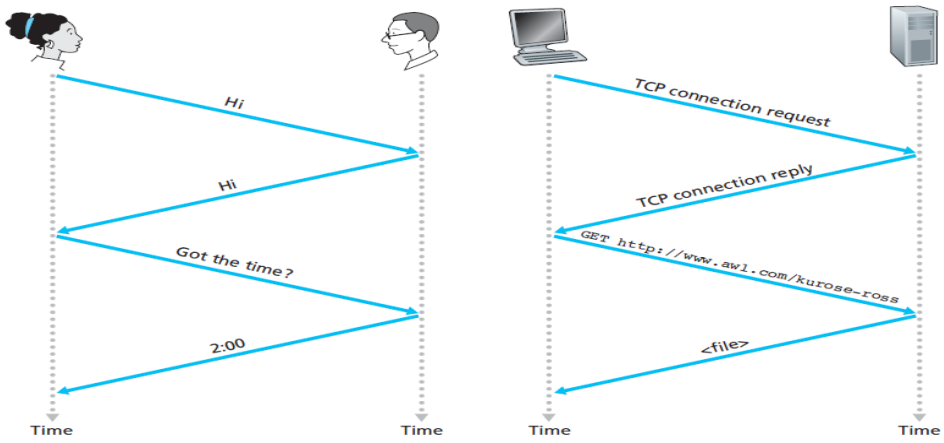

Fig: A human protocol and a computer network protocol

<u>Human protocol</u>

- Entities exchanging messages
- There are specific messages we send, and specific actions we take in response to the received reply messages or other event.

<u>Computer Network Protocols</u>

- Machines rather than humans.
- All communication activity in internet governed by protocols.

 Protocols in routers determine a packet's path from source to destination. Protocols are running everywhere in the Internet. The Internet and computer network make extensive use of protocols. Different protocols are used to accomplish different communication tasks.

2. The Network Edge

- The network edge refers to the area where a device or local network interfaces with the internet. The edge is close to the devices it is communicating with and is the entry point to the network.
- The network edge refers to endpoints. It is the first step between endpoints and the core of the network.
- The internet end systems include desktop computers, servers (web and email servers) and mobile computers (laptops, smart phones, tablets).

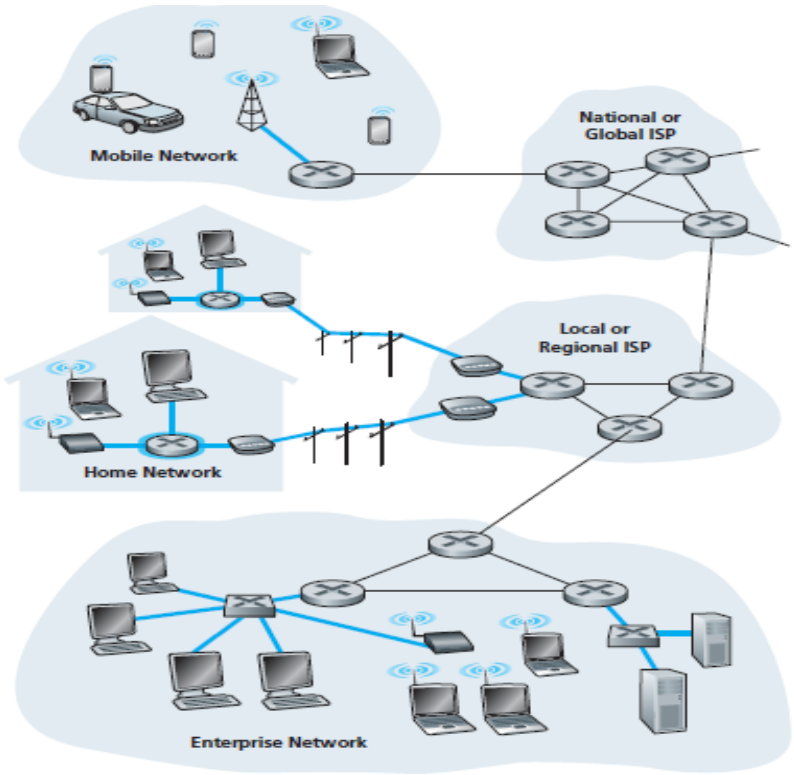

Fig: Access Networks

- End systems are also referred to as *hosts* because they run application programs such as a Web browser program, a Web server program, an e-mail client program, or an e-mail server program.
- Hosts are sometimes further divided into two categories: **clients** and **servers**.

Access Networks

The network that physically connects an end system to the first router (also known as the "edge router") on a path from the end system to any other distant end system.

Home Access: DSL, Cable, FTTH, Dial-Up, and Satellite

i) *DSL (digital subscriber line):* Today, the two most prevalent types of broadband residential access are **digital subscriber line (DSL)** and cable.

- DSL is a wire line transmission technology that transmits data faster.
- A residence typically obtains DSL Internet access from the same local telephone company (telco) that provides its wired local phone access. When DSL is used, a customer's telco is also its ISP.

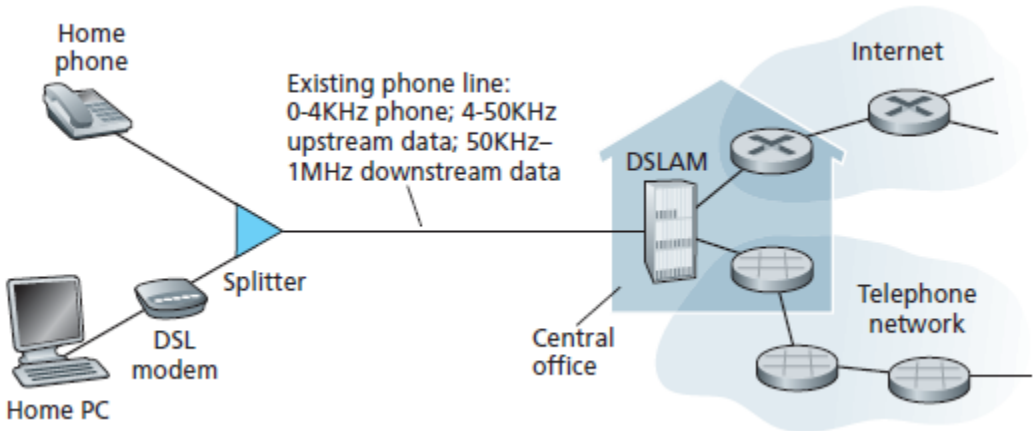

Fig: DSL Internet access

- As shown in Figure, each customer's DSL modem uses the existing telephone line to exchange data with a digital subscriber line access multiplexer (DSLAM) located in the telco's local central office (CO).
- The home's DSL modem takes digital data and translates it to high frequency tones for transmission over telephone wires to the CO; the analog signals from many such houses are translated back into digital format at the DSLAM.
- The residential telephone line carries both data and traditional telephone signals simultaneously.
- On the customer side, a splitter separates the data and telephone signals arriving to the home and forwards the data signal to the DSL modem.
- On the telco side, in the CO, the DSLAM separates the data and phone signals and sends the data into the Internet. Hundreds or even thousands of households connect to a single DSLAM.

ii) Cable: While DSL makes use of the telco's existing local telephone infrastructure, **cable Internet access** makes use of the cable television company's existing cable television infrastructure.
A residence obtains cable Internet access from the same company that provides its cable television.

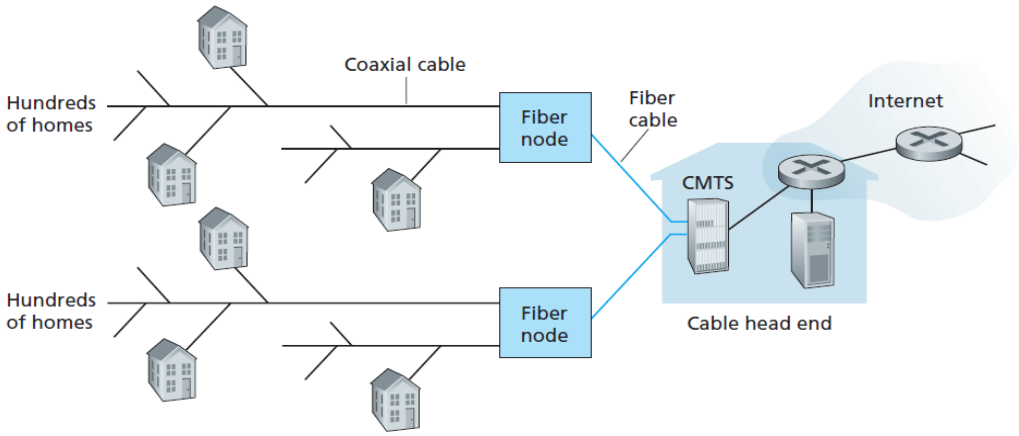

Fig: A hybrid fiber-coaxial access network

- As shows in Figure, fiber optics connects the cable head end to neighborhood-level junctions, from which traditional coaxial cable is then used to reach individual houses and apartments.
- Because both fiber and coaxial cable are employed in this system, it is often referred to as hybrid fiber coax **(HFC).**
- Cable internet access requires special modems, called cable modems. The cable modem is typically an external device and connects to the home PC through an Ethernet port.
- At the cable head end, the cable modem termination system (CMTS) serves a similar function as the DSL network's DSLAM—turning the analog signal sent from the cable modems in many downstream homes back into digital format.
- Cable modems divide the HFC network into two channels, a downstream and an upstream channel.
- One important characteristic of cable Internet access is that it is a shared broadcast medium. In particular, every packet sent by the head end travels downstream on every link to every home and every packet sent by a home travels on the upstream channel to the head end.

iii) FTTH (Fiber To The Home): FTTH includes fiber-optic access solutions designed for residential deployments. In FTTH networks, fibers are directly connected to individual homes or buildings.

- The FTTH can provide an optical fiber path from the CO directly to the home.
- There are two competing optical-distribution network architectures that perform this splitting: active optical networks (AONs) and passive optical networks (PONs).

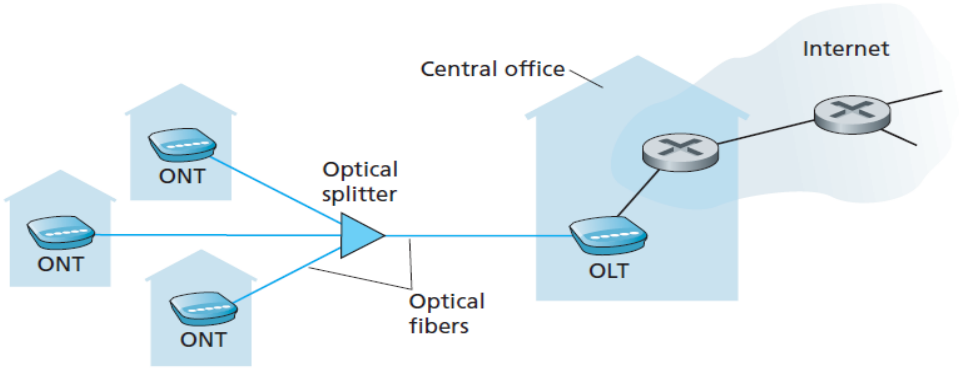

Fig: FTTH Internet access

- Figure shows FTTH using the PON distribution architecture. Each home has an optical network terminator (ONT), which is connected by dedicated optical fiber to a neighborhood splitter.
- The splitter combines a number of homes (typically less than 100) onto a single, shared optical fiber, which connects to an optical line terminator (OLT) in the telco's CO.
- The OLT, providing conversion between optical and electrical signals, connects to the Internet via a telco router.
- In the home, users connect a home router (typically a wireless router) to the ONT and access the Internet via this home router.
- In the PON architecture, all packets sent from OLT to the splitter are replicated at the splitter (similar to a cable head end).
- FTTH can potentially provide Internet access rates in the gigabits per second range.

iv) Dial-Up, and Satellite: *Dial-up* access over traditional phone lines is based on the same model as DSL—a home modem connects over a phone line to a modem in the ISP. Compared with DSL and other broadband access networks, dial-up access is excruciatingly slow at 56 kbps.

Satellite link can be used to connect a residence to the Internet at speeds of more than 1 Mbps; StarBand and HughesNet are two such satellite access providers.

v) Access in the Enterprise (and the Home): Ethernet and WiFi

Ethernet: On corporate and university campuses a local area network (LAN) is used to connect an end system to the edge router.

- There are many types of LAN technologies; Ethernet is the most prevalent access technology in corporate, university, and home networks.
- As shown in Figure, Ethernet users use twisted-pair copper wire to connect to an Ethernet switch.
- With Ethernet access, users typically have 100 Mbps access to the Ethernet switch, whereas servers may have 1 Gbps or even 10 Gbps access.

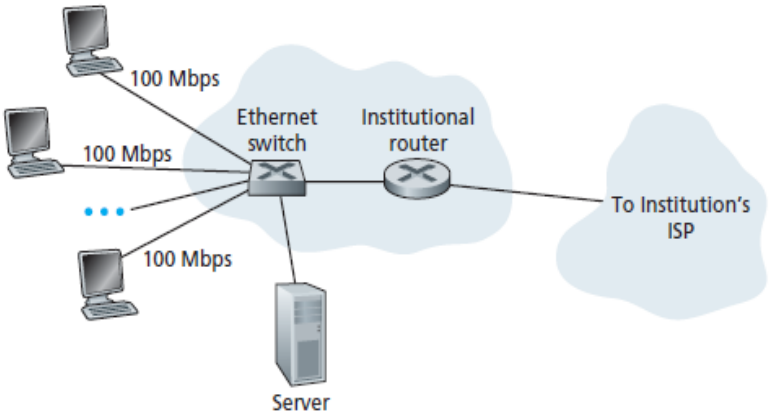

Fig: Ethernet Internet access

WiFi: Increasingly, however, people are accessing the Internet wirelessly from laptops, smart phones, tablets, and other devices.

- In a wireless LAN setting, wireless users transmit/receive packets to/from an access point that is connected into the enterprise's network (most likely including wired Ethernet), which in turn is connected to the wired Internet.
- A wireless LAN user must typically be within a few tens of meters of the access point.
- Wireless LAN access based on IEEE 802.11 technology, more colloquially known as *WiFi*, is now just about everywhere—universities, business offices, cafes, airports, homes, and even in airplanes.
- 802.11 today provides a shared transmission rate of up to 54 Mbps.

3. The Network Core

The network core—the mesh of packet switches and links that interconnects the Internet's end systems. Figure highlights the network core with thick, shaded lines.

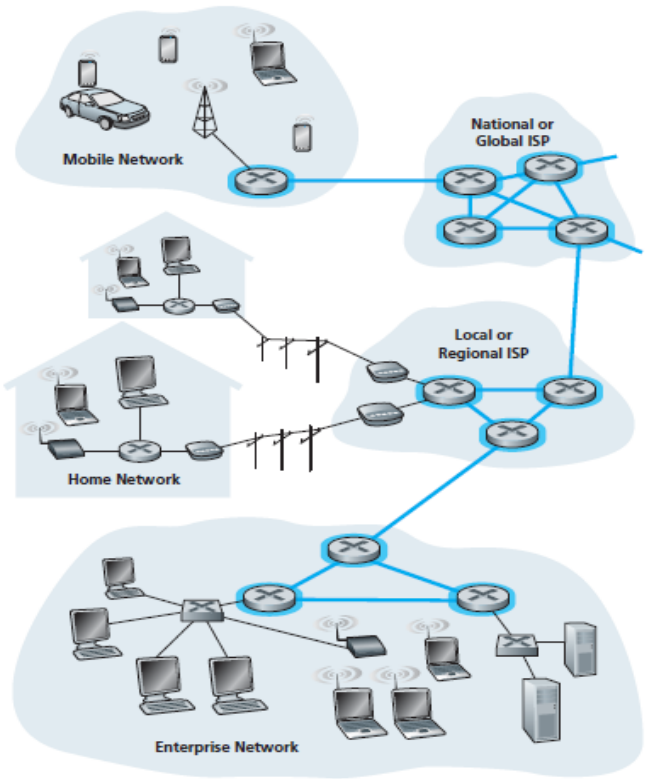

Figure: The network core

- The network edge refers to endpoints. The network core refers to the components that provide services to those at the edge.
- Routing and forwarding are the two key network core functions.

There are two fundamental approaches to moving data through a network of links and switches: **circuit switching** and **packet switching**.

Packet Switching

- In a network application, end systems exchange **messages** with each other. Messages can contain anything the application designer wants.
- Messages may perform a control function or can contain data, such as an email message, a JPEG image, or an MP3audio file.
- To send a message from a source end system to a destination end system, the source breaks long messages into smaller chunks of data known as **packets**.
- Between source and destination, each packet travels through communication links and **packet switches** (for which there are two types, **routers** and **linklayer switches**).
- Packets are transmitted over each communication link at a rate equal to the *full* transmission rate of the link.
- If a source end system or a packet switch is sending a packet of *L* bits over a link with transmission rate *R* bits/sec, then the time to transmit the packet is **L/R seconds**.

Store-and-Forward Transmission: Most packet switches use **store-and-forward transmission** at the inputs to the links.

- Store-and-forward transmission means that the packet switch must receive the entire packet before it can begin to transmit the first bit of the packet onto the outbound link.
- To explore store-and-forward transmission in more detail, consider a simple network consisting of two end systems connected by a single router, as shown in Figure.

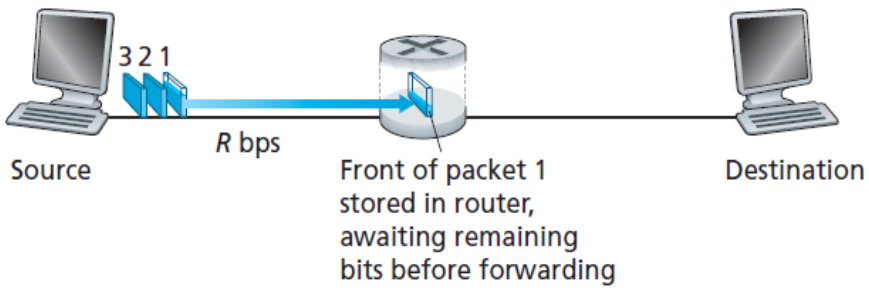

Fig: Store-and-forward packet switching

- A router will typically have many incident links, since its job is to switch an incoming packet onto an outgoing link.
- In this example, the source has three packets, each consisting of *L* bits, to send to the destination.
- The source has transmitted some of packet 1, and the front of packet 1has already arrived at the router. Because the router employs store-and-forwarding, at this instant of time, the router cannot transmit the bits it has received; instead it must first buffer (i.e., "store") the packet's bits.
- Only after the router has received *all* of the packet's bits can it begin to transmit (i.e., "forward") the packet onto the outbound link.

Let's now consider the general case of sending one packet from source to destination over a path consisting of *N* links each of rate *R*. Applying the same logic as above, we see that the end-to-end delay is:

$$d_{\text{end-to-end}} = N\frac{L}{R}$$

 Queuing Delays and Packet Loss: Each packet switch has multiple links attached to it. For each attached link, the packet switch has an **output buffer** (also called an **output queue**), which stores packets that the router is about to send into that link.

- The output buffers play a key role in packet switching. If an arriving packet needs to be transmitted onto a link but finds the link busy with the transmission of another packet, the arriving packet must wait in the output buffer.
- Thus, in addition to the store-and-forward delays, packets suffer output buffer **queuing delays**. These delays are variable and depend on the level of congestion in the network.
- The amount of buffer space is finite, an arriving packet may find that the buffer is completely full with other packets waiting for transmission. In this case, **packet loss** will occur—either the arriving packet or one of the already-queued packets will be dropped.

Figure illustrates a simple packet-switched network.

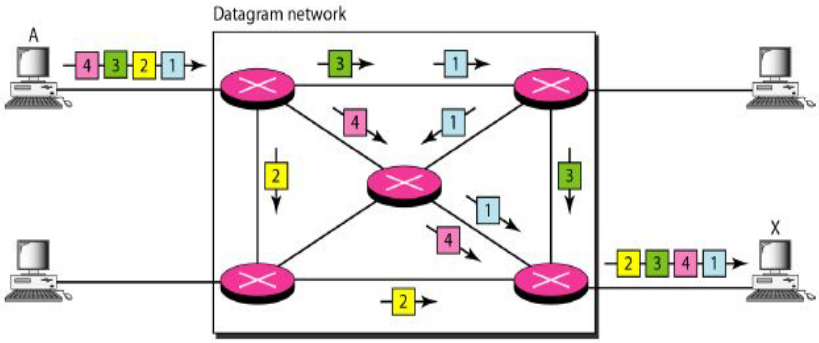

Fig: Packet switching

 Forwarding Tables and Routing Protocols: In the Internet, every end system has an address called an IP address. When a source end system wants to send a packet to a destination end system, the source includes the destination's IP address in the packet's header.

- When a packet arrives at a router in the network, the router examines a portion of the packet's destination address and forwards the packet to an adjacent router.
- More specifically, each router has a **forwarding table** that maps destination addresses (or portions of the destination addresses) to that router's outbound links.
- When a packet arrives at a router, the router examines the address and searches its forwarding table, using this destination address, to find the appropriate outbound link. The router then directs the packet to this outbound link.
- The Internet has a number of special **routing protocols** that are used to automatically set the forwarding tables. A routing protocol may, for example, determine the shortest path from each router to each destination and use the shortest path results to configure the forwarding tables in the routers.

Circuit Switching

- In circuit-switched networks, the resources needed along a path (buffers, link transmission rate) to provide for communication between the end systems are *reserved* for the duration of the communication session between the end systems.
- In packet-switched networks, these resources are *not* reserved; a session's messages use the resources on demand, and as a consequence, may have to wait (that is, queue) for access to a communication link.
- Traditional telephone networks are examples of circuit-switched networks.
- A circuit-switched network is made of a set of switches connected by physical

links, in which each link is divided into n channels.

- Each link is divided into n (n is 3 in the figure) channels by using FDM or TDM.

 Three phases: A circuit-switched network consists of ***3 phases***: 1) Setup phase (establish),2) Data transfer phase (transfer), 3) Tear down phase (disconnect).

- **Setup Phase**: Before the two parties can communicate, a dedicated circuit needs to be established. The end systems are normally connected through dedicated lines to the switches, so connection setup means creating dedicated channels between the switches.
- **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.

Figure illustrates a circuit-switched network. In this network, the four circuit switches are interconnected by four links. Each of these links has four circuits, so that each link can support four simultaneous connections.

- The hosts are each directly connected to one of the switches. When two hosts want to communicate, the network establishes a dedicated **end-to-end connection** between the two hosts.

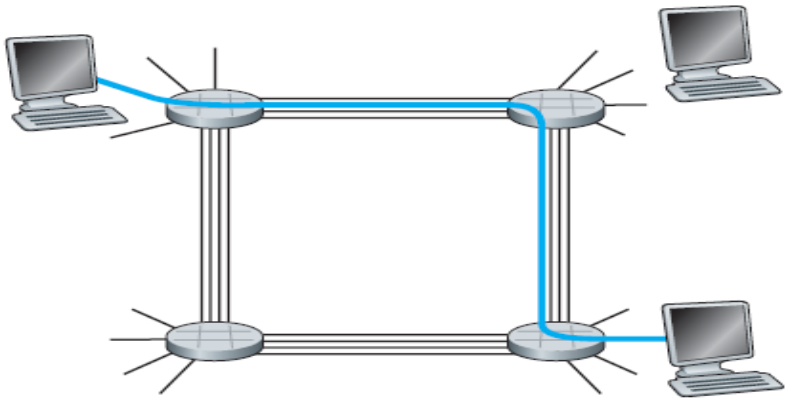

Fig: A simple circuit-switched network consisting of four switches and four links

For example: When end system A needs to communicate with end system B, system A needs to request a connection to B that must be accepted by all switches as well as by B itself. This is called the ***setup phase***; after the dedicated path made of connected circuits (channels) is established, the ***data-transfer*** phase can take place. After all data have been transferred, the circuits are **tearing down.**
In circuit switching, the resources need to be reserved during the setup phase; the resources remain dedicated for the entire duration of data transfer until the teardown phase.

 Multiplexing in Circuit-Switched Networks: A circuit in a link is implemented with either **frequency-division multiplexing (FDM)** or **time-division multiplexing (TDM)**.

With ***FDM***, the frequency spectrum of a link is divided up among the connections established across the link.

- The link dedicates a frequency band to each connection for the duration of the connection. The width of the band is called, not surprisingly, the **bandwidth**.
- FM radio stations also use FDM to share the frequency spectrum between 88 MHz and 108 MHz, with each station being allocated a specific frequency band.

For a ***TDM*** link, time is divided into frames of fixed duration, and each frame is divided into a fixed number of time slots. When the network establishes a

connection across a link, the network dedicates one time slot in every frame to this connection.

Figure shows FDM and TDM for a specific network link supporting up to four circuits. For FDM, the frequency domain is segmented into four bands, each of bandwidth 4 kHz. For TDM, the time domain is segmented into frames, with four time slots in each frame

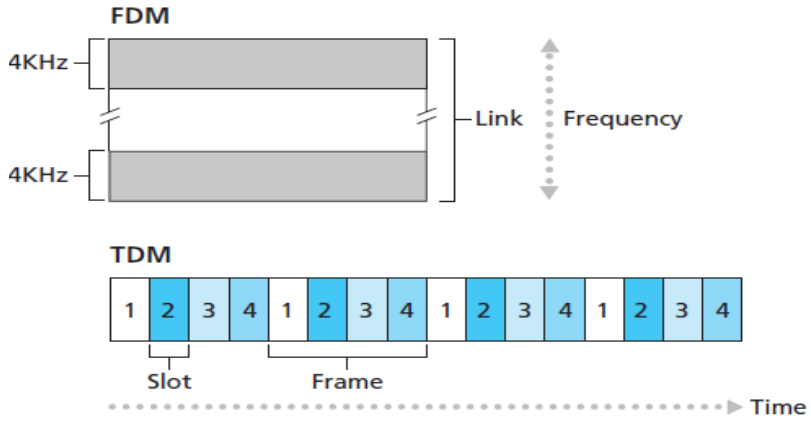

Fig: FDM & TDM

4. Delay, Loss, and Throughput in Packet-Switched Networks

Overview of Delay in Packet-Switched Networks

A packet starts in a host (the source), passes through a series of routers, and ends its journey in another host (the destination). As a packet travels from one node (host or router) to the subsequent node (host or router) along this path, the packet suffers from several types of delays at *each* node along the path.

The most important of these delays are the **nodal processing delay, queuing delay, transmission delay,** and **propagation delay**; together, these delays accumulate to give a **total nodal delay**.

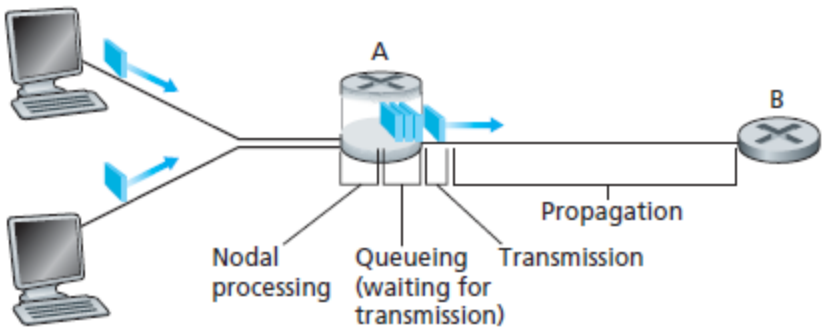

Fig: The nodal delay at router A

Types of Delay: The end-to-end route between source and destination, a packet is sent from the upstream node through router A to router B. Our goal is to characterize the nodal delay at router A.

i) Processing Delay: The time required to examine the packet's header and determine where to direct the packet is part of the **processing delay**.

- The processing delay can also include other factors, such as the time needed to check for bit-level errors in the packet that occurred in transmitting the packet's bits from the upstream node to router A.
- Processing delays in high-speed routers are typically on the order of microseconds or less. After this nodal processing, the router directs the packet to the queue that precedes the link to router B.

ii) **Queuing Delay:** At the queue, the packet experiences a **queuing delay** as it waits to be transmitted onto the link.

- The length of the queuing delay of a specific packet will depend on the number of earlier-arriving packets that are queued and waiting for transmission onto the link.
- If the queue is empty and no other packet is currently being transmitted, then our packet's queuing delay will be zero.

On the other hand, if the traffic is heavy and many other packets are also waiting to be transmitted, the queuing delay will be long.

- Queuing delays can be on the order of microseconds to milliseconds in practice.

iii) **Transmission Delay:** This is the amount of time required to push (that is, transmit) all of the packet's bits into the link.

- Denote the length of the packet by *L* bits, and denote the transmission rate of the link from router A to router B by *R* bits/sec.
- For example, for a 10 Mbps Ethernet link, the rate is *R* = 10 Mbps; for a 100 Mbps Ethernet link, the rate is *R* = 100 Mbps.
- The **transmission delay** is *L/R.* Transmission delays are typically on the order of microseconds to milliseconds in practice.

iv) **Propagation Delay:** Once a bit is pushed into the link, it needs to propagate to router B. The time required to propagate from the beginning of the link to router B is the **propagation delay**.

- The bit propagates at the propagation speed of the link. The propagation speed depends on the physical medium of the link (that is, fiber optics, twisted-pair copper wire, and so on)
- The propagation delay is the distance between two routers divided by the propagation speed. That is, the propagation delay is *d/s,* where *d* is the distance between router A and router B and *s* is the propagation speed of the link. Propagation delays are on the order of milliseconds.

The total ***nodal delay*** is given by

$$d_{nodal} = d_{proc} + d_{queue} + d_{trans} + d_{prop}$$

Queuing Delay and Packet Loss

The queuing delay can vary from packet to packet. For example, if 10 packets arrive at an empty queue at the same time, the first packet transmitted will suffer no queuing delay, while the last packet transmitted will suffer a relatively large queuing delay (while it waits for the other nine packets to be transmitted).

On the other hand, if the traffic is heavy and many other packets are also waiting to be transmitted, the queuing delay will be long.

The ratio *of* the **traffic intensity is** ***La/R***

Let ***a*** denote the average rate at which packets arrive to the queue (*a* is units of packets/sec), ***R*** is the transmission rate, i.e., it is the rate (in bits/sec) and **L** is average packet length (in bits).

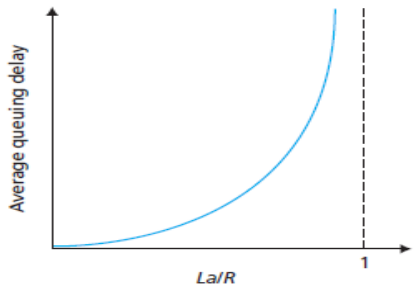

Fig: Dependence of average queuing delay on traffic intensity

The fact that as the traffic intensity approaches 1, the average queuing delay increases rapidly. A small percentage increase in the intensity will result in a much larger percentage-wise increase in delay.

Packet Loss: we have assumed that the queue is capable of holding an infinite number of packets. In reality a queue preceding a link has finite capacity, although the queuing capacity greatly depends on the router design and cost. Because the queue capacity is finite, packet delays do not really approach infinity as the traffic intensity approaches 1. Instead, a packet can arrive to find a full queue.

With no place to store such a packet, a router will **drop** that packet; that is, the packet will be **lost**.

- **End-to-End Delay:** nodal delay is the delay at a single router. Let's now consider the total delay from source to destination. To get a handle on this concept, suppose there are *N*-1 routers between the source host and the destination host.

The nodal delays accumulate and give an end-to-end delay,

$$d_{\text{end-end}} = N\,(d_{\text{proc}} + d_{\text{trans}} + d_{\text{prop}})$$

Throughput in Computer Networks

In data transmission, network throughput is **the amount of data moved successfully from one place to another in a given time period**, and typically measured in bits per second (bps), as in megabits per second (Mbps) or gigabits per second (Gbps).

Network throughput **refers to how much data can be transferred from source to destination within a given timeframe.**

To define throughput, consider transferring a large file from Host A to Host B across a computer network.

The **instantaneous throughput** at any instant of time is the rate (in bits/sec) at which Host B is receiving the file.

If the file consists of *F* bits and the transfer takes *T* seconds for Host B to receive all *F* bits, then the **average throughput** of the file transfer is *F/T* bits/sec.

Example: Figure (a) shows two end systems, a server and a client, connected by two communication links and a router. Consider the throughput for a file transfer from the server to the client.

Let *Rs* denote the rate of the link between the server and the router; and *Rc* denote the rate of the link between the router and the client.

For this simple two-link network, the throughput is min{*Rc*, *Rs*}, that is, it is the transmission rate of the **bottleneck link**. Having determined the throughput, we can now approximate the time it takes to transfer a large file of *F* bits from server to client as *F*/min{*Rs*, *Rc*}.

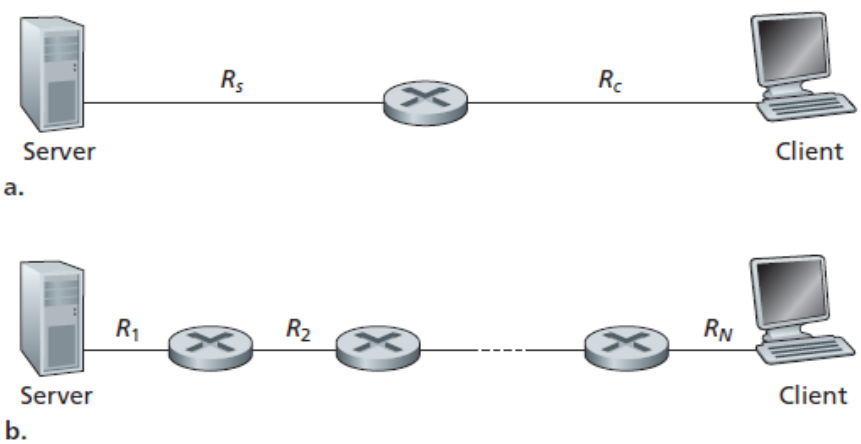

Fig: Throughput for a file transfer from server to client

Figure (b) now shows a network with *N* links between the server and the client, with the transmission rates of the *N* links being *R*1, *R*2,..., *RN*. Applying the same analysis as for the two-link network, we find that the throughput for a file transfer from server to client is min{*R*1, *R*2,..., *RN*},

5. Reference Models

Layered Architecture or Protocol Layering

A layered architecture allows us to discuss a well-defined, specific part of a large and complex system. In layered architecture of network model, one whole network process is divided into small tasks. Each small task is then assigned to a particular layer which works dedicatedly to process the task only. Every layer does only specific work.
The layer provides the same service to the layer above it, and uses the same services from the below it.

In data communication and networking, a protocol defines the rules that both the sender and receiver and all intermediate devices need to follow to be able to communicate effectively. When communication is simple, we may need only one simple protocol; when the communication is complex, we may need to divide the task between different layers, in which case we need a protocol at each layer, or protocol layering.

Scenarios

Let us develop two simple scenarios to better understand the need for protocol layering.
First Scenario: In the first scenario, communication is so simple that it can occur in only one layer. Assume Maria and Ann are neighbors with a lot of common ideas. Communication between Maria and Ann takes place in one layer.

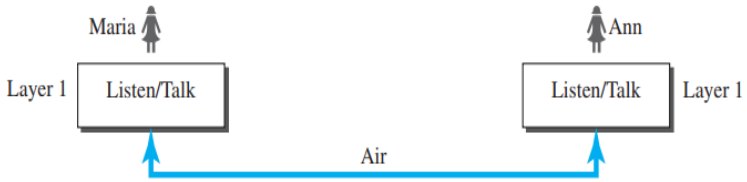

Fig: A single-layer protocol

Second Scenario: In the second scenario, we assume that Ann is offered a higher-level position in her company, but needs to move to another branch located in a city very far from Maria. The two friends still want to continue their communication and exchange ideas by using protocol layering.

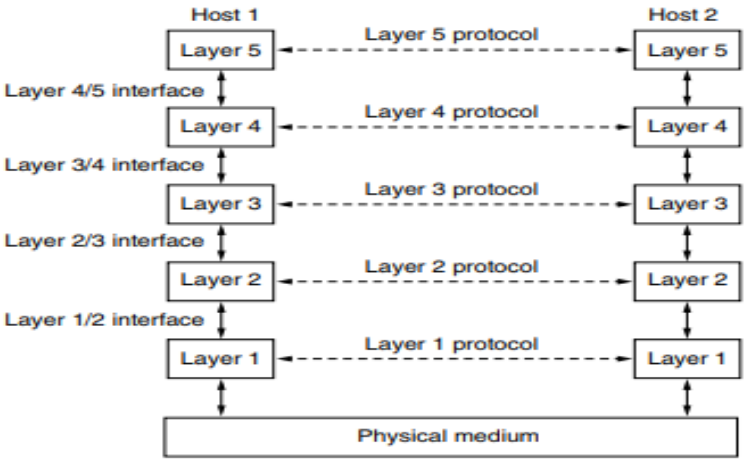

Fig: multiple protocols layering

Two models have been devised to define computer network operations: the ***TCP/IP protocol suite*** and the ***OSI model.*** The protocol layering is used in both models.

Reference Models

i) OSI model

The OSI model is based on a proposal developed by the International Standards Organization (ISO). The model is called the ISO OSI (Open Systems Interconnection), which allows different systems to communicate.

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 a layered framework for the design of network systems that allows communication between all types of computer systems.

It consists of ***seven layers***: 1. Physical Layer, 2. Data link Layer, 3. Network Layer, 4.transport Layer, 5. Session Layer, 6. Presentation layer, 7. Application Layer.

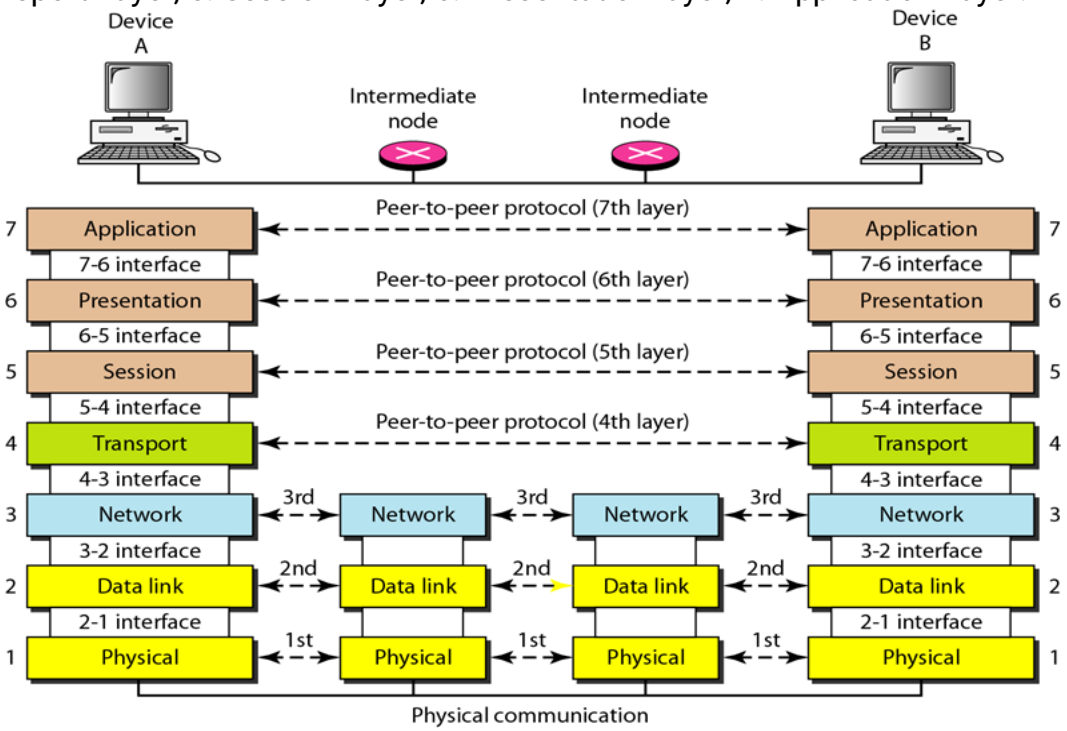

Fig: The interaction between layers in the OSI model

i) **Physical layer:** the physical layer is responsible for movement of individual bits from one node to the next.

- The physical layer required to carry a bit stream over a physical medium.
- It deals with the mechanical and electrical specifications of the interface and transmission medium.

Responsibilities of physical layer:

Physical characteristics of interfaces and medium: The physical layer defines the characteristics of the interface between the devices and the transmission medium. It also defines the type of transmission medium.

Physical topology: The physical topology defines how devices are connected to make a network.

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

ii) **Data Link Layer:** The data link layer is responsible for moving frames from one node to the next.

Frame: Frame is a series of bits that form a unit of data.

Responsibilities of the data link layer:

Framing: The data link layer divides the stream of bits received from the network layer into manageable data units called frames.

Physical addressing: 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.

Flow control: If the rate at which the data are absorbed by the receiver is less than the rate at which data are produced in the sender, the data link layer imposes a flow control mechanism to avoid overwhelming the receiver.

Error control: The data link layer adding a mechanisms to detect and retransmit damaged or lost frames. It also uses a mechanism to recognize duplicate frames.

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.

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

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

Responsibilities of the network layer

Logical addressing: Addressing system to help to differentiate the source and destination systems. The network layer adds a header to the packet coming from the upper layer that includes the logical addresses of the sender and receiver.

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

iv) **Transport Layer**: The transport layer is responsible for process-to-process delivery of the entire message. A process is an application program running on a host.

Responsibilities of the transport layer:

Service-point addressing: 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 includes a type of address called a service-point address (or port

address).
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.
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.

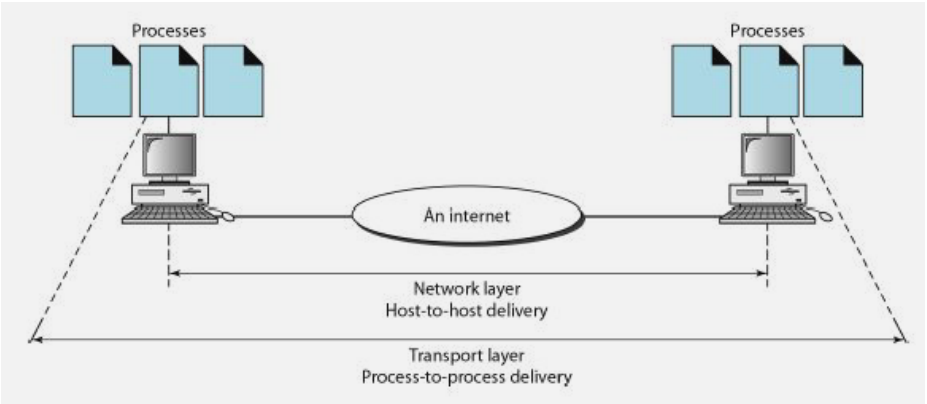

Fig: process-to-process delivery of a message

v) Session Layer: The session layer is responsible for dialog control and synchronization. The session layer is the network dialog controller. It establishes, maintains, and synchronizes the interaction among communicating systems
Responsibilities of the session layer:
Dialog control: The session layer allows two systems to enter into a dialog. It allows the communication between two processes to take place in either half-duplex (one way at a time) or full-duplex (two ways at a time) mode.
Synchronization: The session layer allows a process to add checkpoints, or synchronization points, to a stream of data.
For example, if a system is sending a file of 2000 pages, it is advisable to insert checkpoints after every 100 pages to ensure that each 100-page unit is received and acknowledged independently. In this case, if a crash happens during the transmission of page 523, the only pages that need to be resent after system recovery are pages 501 to 523. Pages previous to 501 need not be resent.

vi) Presentation Layer: The presentation layer is responsible for translation, compression, and encryption. The presentation layer is concerned with the syntax and semantics of the information exchanged between two systems.
Responsibilities of the presentation layer:
Translation: The presentation layer at the sender machine changes the information from its sender-dependent format into a common format. The presentation layer at the receiving machine changes the common format into its receiver-dependent format.
Encryption: To carry sensitive information, a system must be able to ensure privacy. Encryption means that the sender transforms the original information to another form and sends the resulting message out over the network. Decryption reverses the original process to transform the message back to its original form.
Compression: Data compression reduces the number of bits contained in the information. Data compression becomes particularly important in the transmission of multimedia such as text, audio, and video.

vii) Application Layer: The application layer is responsible for providing services to

the user.
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.

Services provided by the application layer:

Network virtual terminal: A network virtual terminal is a software version of a physical terminal, and it allows a user to log on to a remote host.

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.

ii) TCP/IP Protocol Suite

The TCPIIP protocol suite was developed prior to the OSI model. The original TCP/IP protocol suite was defined as having ***four layers***: Host-To-Network (Network Interface), Internet, Transport, and Application.

The Host-To-Network layer is equivalent to the combination of the physical and data link layers. The internet layer is equivalent to the network layer, and the application layer is roughly doing the job of the session, presentation, and application layers.

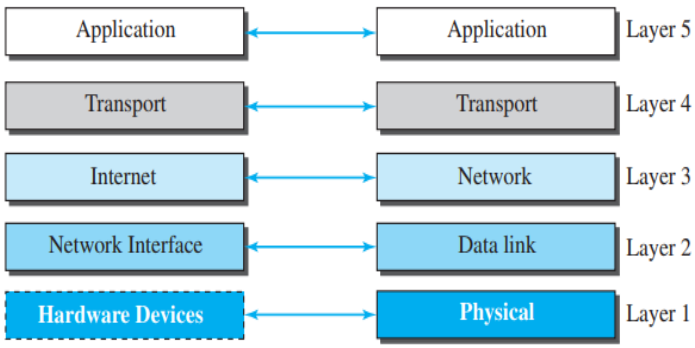

Fig: Layers in the TCP/IP protocol suite

- **Physical Layer:** We can say that the physical layer is responsible for carrying individual bits in a frame across the link. The physical layer is the lowest level in the TCP/IP protocol suite. There is a hidden layer, the transmission media, under the physical layer. Two devices are connected by a transmission medium (cable or air).The transmission medium does not carry bits; it carries electrical or optical signals.

- **Data-link Layer:** the data-link layer is responsible for taking the datagram and moving it across the link. The link can be a wired LAN with a link-layer switch, a wireless LAN, a wired WAN, or a wireless WAN. We can also have different protocols used with any link type.

TCP/IP does not define any specific protocol for the data-link layer, but it uses the HDLC and PPP protocols. It supports all the standard and proprietary protocols.

- **Network Layer:** The network layer is responsible for creating a connection between the source computer and the destination computer. The communication at the network layer is host-to-host. There can be several routers from the source to the destination; the routers in the path are responsible for choosing the best route for each packet.

In network layer the *main protocol* is **Internet Protocol** (IP), which defines the format of the packet, called a datagram at the network layer. IP also defines the format and the structure of addresses used in this layer.
The network layer also has some auxiliary protocols that help IP in its delivery and routing tasks. The **Internet Control Message Protocol** (ICMP) helps IP to report some problems when routing a packet. The **Internet Group Management Protocol** (IGMP) is another protocol that helps IP in multitasking. The **Dynamic Host Configuration Protocol** (DHCP) helps IP to get the network-layer address for a host. The **Address Resolution Protocol** (ARP) is a protocol that helps IP to find the link-layer address of a host or a router when its network-layer address is given.

- **Transport Layer:** The logical connection at the transport layer is also end-to-end. The transport layer at the source host gets the message from the application layer, encapsulates it in a transport layer packet called a segment or a user datagram.
The main protocol, **Transmission Control Protocol** (TCP), is a connection-oriented protocol that first establishes a logical connection between transport layers at two hosts before transferring data. It creates a logical pipe between two TCPs for transferring a stream of bytes.
The other common protocol, **User Datagram Protocol** (UDP), is a connectionless protocol that transmits user datagram without first creating a logical connection. In UDP, each user datagram is an independent entity without being related to the previous or the next one.

- **Application Layer:** The logical connection between the two application layers is end to-end. The two application layers exchange *messages* between each other as though there were a bridge between the two layers.
The **Hypertext Transfer Protocol** (HTTP) is a vehicle for accessing the World Wide Web (WWW). The **Simple Mail Transfer Protocol** (SMTP) is the main protocol used in electronic mail (e-mail) service. The **File Transfer Protocol** (FTP) is used for transferring files from one host to another. The **Terminal Network** (TELNET) and **Secure Shell** (SSH) are used for accessing a site remotely. The **Simple Network Management Protocol** (SNMP) is used by an administrator to manage the Internet at global and local levels. The **Domain Name System** (DNS) is used by other protocols to find the network-layer address of a computer.

6. Transmission Media

Introduction: Transmission media are actually located below the physical layer and are directly controlled by the physical layer. We could say that transmission media belong to layer zero. Figure shows the position of transmission media in relation to the physical layer.

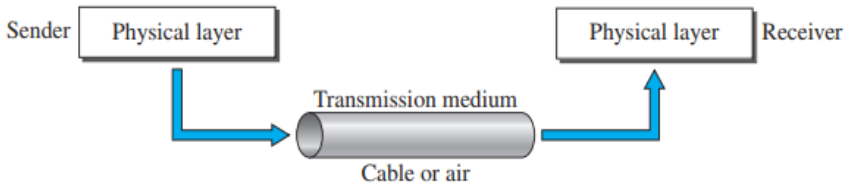

Fig: Transmission medium and physical layer

A transmission medium can be broadly defined as anything that can carry information from a source to a destination. It is also called physical medium.
The transmission medium is usually free space, metallic cable, or fiber-optic cable. The information is usually a signal that is the result of a conversion of data from another form.

Transmission media can be divided into **two broad categories**: guided and unguided.

Guided media include twisted-pair cable, coaxial cable, and fiber-optic cable. Unguided medium is free space.

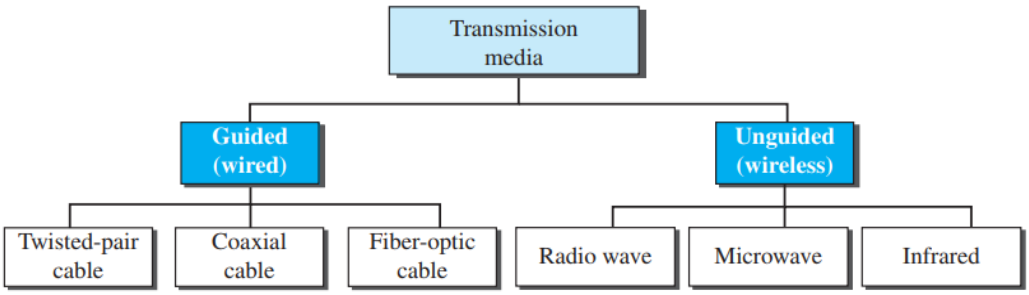

Fig: Classes of transmission media

i) Guided Media: Wired

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.

 Twisted-Pair Cable: A twisted pair consists of two conductors (normally copper), each with its own plastic insulation, twisted together, as shown in Figure

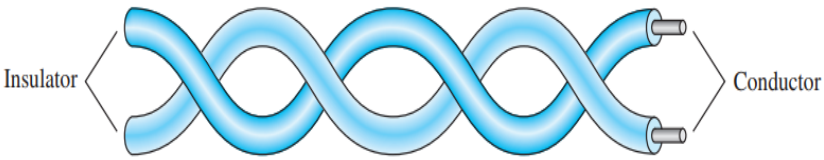

Fig: Twisted-pair cable

One of the wires is used to carry signals to the receiver, and the other is used only as a ground reference. The receiver uses the difference between the two. In addition to the signal sent by the sender on one of the wires, interference (noise) and crosstalk may affect both wires and create unwanted signals.

Unshielded Versus Shielded Twisted-Pair Cable: The most common twisted-pair cable used in communications is referred to as unshielded twisted-pair (UTP).

IBM has also produced a version of twisted-pair cable for its use, called shielded twisted-pair (STP). STP cable has a metal foil or braided mesh covering that encases each pair of insulated conductors. Although metal casing improves the quality of cable by preventing the penetration of noise or crosstalk, it is bulkier and more expensive.

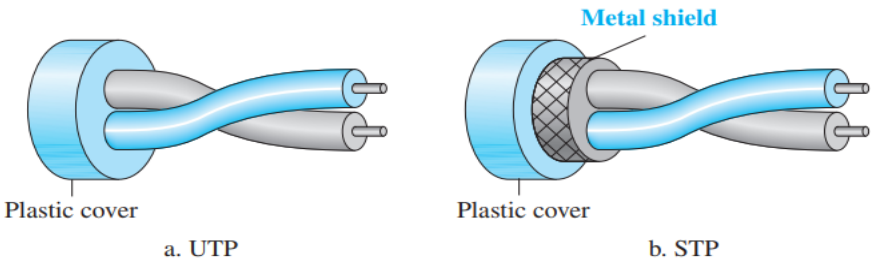

Fig: UTP and STP cables

Performance: One way to measure the performance of twisted-pair cable is to compare attenuation versus frequency and distance. A twisted-pair cable can pass a wide range of frequencies.

Applications: Twisted-pair cables are used in telephone lines to provide voice and data channels.

- Local-area networks also use twisted-pair cables.

 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. Instead of having two wires, coax has a central core conductor of solid or stranded wire (usually copper) enclosed in an insulating sheath, which is, in turn, encased in an outer conductor of metal foil, braid, or a combination of the two.
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.

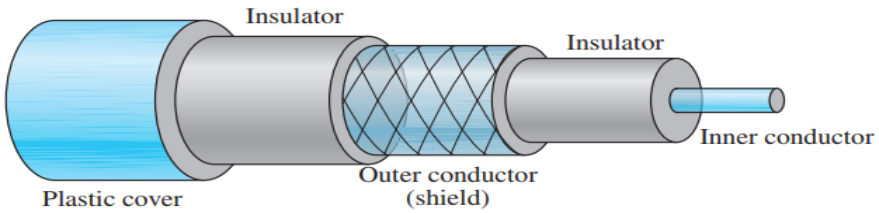

Fig: Coaxial cable

Performance we measure the performance of a coaxial cable. The attenuation is much higher in coaxial cable than in twisted-pair cable. In other words, although coaxial cable has a much higher bandwidth, the signal weakens rapidly and requires the frequent use of repeaters.

Applications Coaxial cable was widely used in analog telephone networks where a single coaxial network could carry 10,000 voice signals.

- Later it was used in digital telephone networks where a single coaxial cable could carry digital data up to 600 Mbps.
- However, coaxial cable in telephone networks has largely been replaced today with fiber optic cable.
- Cable TV networks also use coaxial cables. Later, however, cable TV providers replaced most of the media with fiber-optic cable.
- Another common application of coaxial cable is in traditional Ethernet LANs. Because of its high bandwidth, and consequently high data rate, coaxial cable was chosen for digital transmission in early Ethernet LANs.

 Fiber-Optic Cable: A fiber-optic cable is made of glass or plastic and transmits signals in the form of light.
Light travels in a straight line as long as it is moving through a single uniform substance. If a ray of light traveling through one substance suddenly enters another substance (density), the ray changes direction.

- Figure shows how a ray of light changes direction when going from a more dense to a less dense substance. As the figure shows, if the angle of incidence I 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.

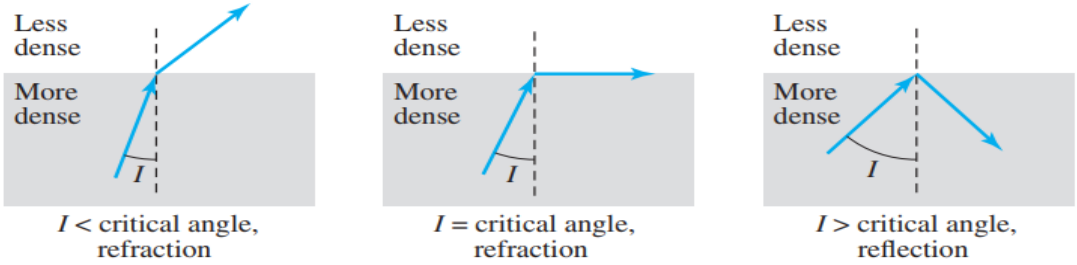

Fig: Bending of light ray

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.

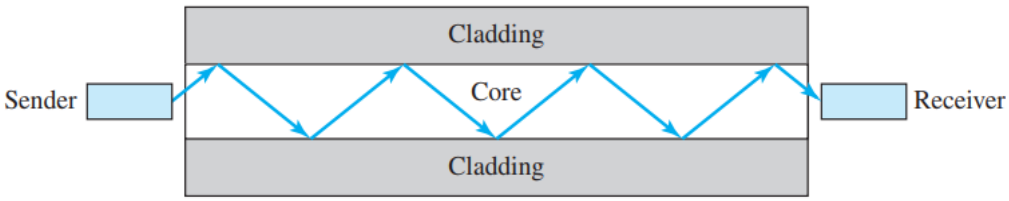

Fig: Optical fiber

Cable Composition: Figure shows the composition of a typical fiber-optic cable. The outer jacket is made of Teflon. Inside the jacket are Kevlar strands to strengthen the cable. Below the Kevlar is another plastic coating to cushion the fiber. The fiber is at the center of the cable, and it consists of cladding and core.

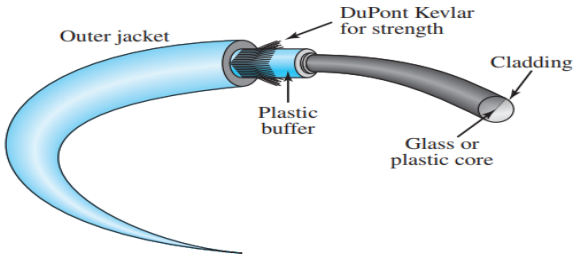

Fig: Fiber construction

Performance: Attenuation is flatter than in the case of twisted-pair cable and coaxial cable. The performance is such that we need fewer repeaters when we use fiber-optic cable.

Advantages and Disadvantages of Optical Fiber:
Advantages: Fiber-optic cable has several advantages over metallic cable (twisted-pair or coaxial).

- Higher bandwidth.
- Less signal attenuation..
- Immunity to electromagnetic interference.
- Resistance to corrosive materials.
- Light weight.

.
Disadvantages There are some disadvantages in the use of optical fiber.

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

ii) Unguided Media: Wireless

Unguided medium 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.
Figure shows the part of the electromagnetic spectrum, ranging from 3 kHz to 900 THz, used for wireless communication.

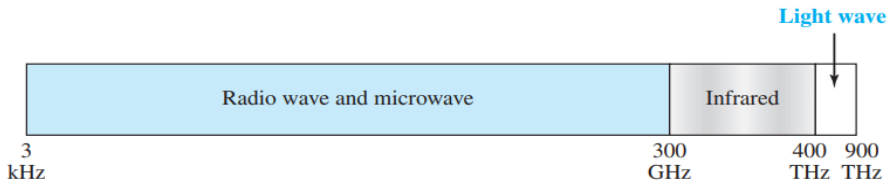

Fig: Electromagnetic spectrum for wireless communication

Unguided signals can travel from the source to the destination in several ways: ground propagation, sky propagation, and line-of-sight propagation, as shown in Figure 7.18.

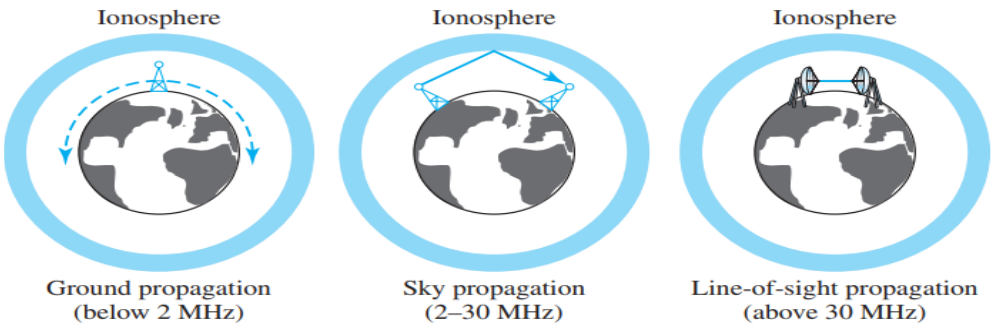

Fig: Propagation methods

In ***ground propagation***, radio waves travel through the lowest portion of the atmosphere, hugging the earth. These low-frequency signals emanate in all directions from the transmitting antenna and follow the curvature of the planet.
In ***sky propagation***, higher-frequency radio waves radiate upward into the ionosphere (the layer of atmosphere where particles exist as ions) where they are reflected back to earth. This type of transmission allows for greater distances with lower output power.
In ***line-of-sight propagation***, very high-frequency signals are transmitted in straight lines directly from antenna to antenna.

 Radio Waves: The electromagnetic waves ranging in frequencies between 3 kHz and 1 GHz are normally called radio waves.

- Radio waves are omnidirectional. When an antenna transmits radio waves, they are propagated in all directions. A sending antenna sends waves that can be received by any receiving antenna.
- Radio waves, particularly those waves that propagate in the sky mode, can travel long distances. This makes radio waves a good candidate for long-distance broadcasting such as AM radio.
- Radio waves, particularly those of low and medium frequencies, can penetrate walls. It is an advantage because, for example, an AM radio can receive signals inside a building.

Omnidirectional Antenna Radio waves use omnidirectional antennas that send out signals in all directions. Figure shows an omnidirectional antenna.

Fig: Omnidirectional antenna

Applications The omnidirectional characteristics of radio waves make them useful for multicasting, in which there is one sender but many receivers.
AM and FM radio, television, cordlessphones, and paging are examples of multicasting.

 Microwaves Electromagnetic waves having frequencies between 1 and 300 GHz are called microwaves. Microwaves are unidirectional. This means that the sending and receiving antennas need to be aligned.
The following describes some ***characteristics*** of microwave propagation:

- Microwave propagation is line-of-sight. Since the towers with the mounted antennas need to be in direct sight of each other, towers that are far apart need to be very tall.
- Very high-frequency microwaves cannot penetrate walls. This characteristic can be a disadvantage if receivers are inside buildings.

Unidirectional Antenna 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.

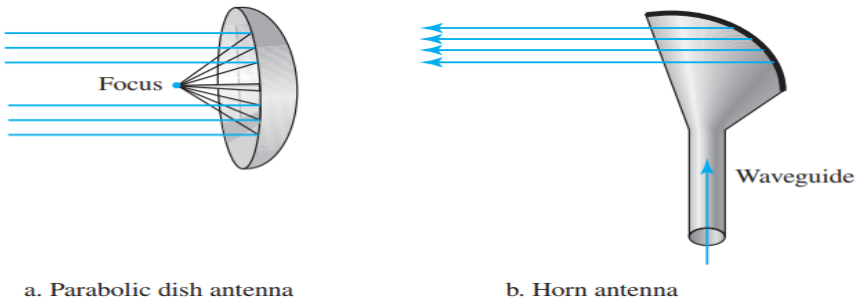

Fig: Unidirectional antennas

Applications: Microwaves, due to their unidirectional properties, are very useful when unicast (oneto-one) communication is needed between the sender and the receiver.
They are used in cellular phones, satellite networks, and wireless LANs.
 Infrared: Infrared waves, with frequencies from 300 GHz to 400 THz. It 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.
Infrared signals can be used for short-range communication in a closed area using line-of-sight propagation

7. Example Networks

i) Internet

A Brief History: A network is a group of connected communicating devices such as computers and printers. The Internet collaboration of more than hundreds of thousands of interconnected networks.
Private individuals as well as various organizations such as government agencies, schools, research facilities, corporations, and libraries in more than 100 countries use the Internet. Millions of people are users. The extraordinary communication system only came into being in 1969.

In the *mid-1960s*, Computers from different manufacturers were unable to communicate with one another. The Advanced Research Projects Agency (ARPA) in the Department of Defense (DoD) was interested in finding a way to connect computers.

In 1967, at an Association for Computing Machinery (ACM) meeting, ARPA presented its ideas for ARPANET, a small network of connected computers. The idea was that each host computer would be attached to a specialized computer, called an interface message processor (IMP).

By *1969*, ARPANET was a reality. Four nodes, at the Universities were connected via the IMPs to form a network. Software called the Network Control Protocol (NCP) provided communication between the hosts.

In *1972*, Vint Cerf and Bob Kahn, both of whom were part of the core ARPANET group, collaborated on what they called the Internetting Project.

Cerf and Kahn's landmark *1973* paper outlined the protocols to achieve end-to-end delivery of packets. This paper on Transmission Control Protocol (TCP) included concepts such as encapsulation, the datagram, and the functions of a gateway.
After that split TCP into two protocols: Transmission Control Protocol (TCP) and Internetworking Protocol (IP).

The Internet Today: 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.

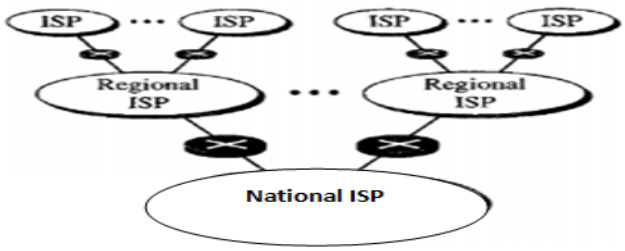

Fig: Structure of a national ISP

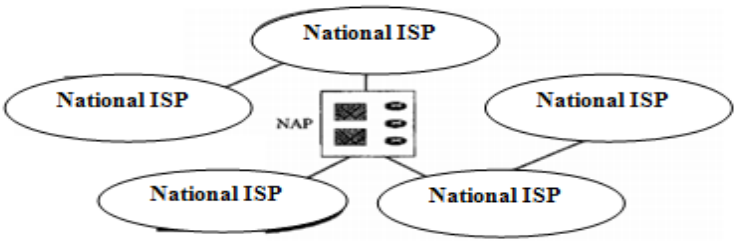

Fig: Interconnection of national ISPs

ii) Third-Generation Mobile Phone Networks

People love to talk on the phone even more than they like to surf the Internet, and this has made the mobile phone network the most successful network in the world. It has more than four billion subscribers worldwide.
First-generation mobile phone systems transmitted voice calls as continuously varying (analog) signals rather than sequences of (digital) bits. **AMPS** (**Advanced Mobile Phone System**), which was deployed in the United States in 1982, was a widely used first generation system.
Second-generation mobile phone systems switched to transmitting voice calls in digital form to increase capacity, improve security, and offer text messaging. **GSM** (**Global System for Mobile communications**), which was deployed starting in 1991 and has become the most widely used mobile phone system in the world, is a 2G system.
The *third generation*, or 3G, systems were initially deployed in 2001 and offer both digital voice and broadband digital data services. **UMTS** (**Universal Mobile Telecommunications System**), also called **WCDMA** (**Wideband Code Division Multiple Access**), is the main 3G system that is being rapidly deployed worldwide.

iii) Wireless LANs: 802.11

The wireless LAN standard was dubbed 802.11. A common slang name for it is **WiFi** but it is an important standard and deserves respect, so we will call it by its proper name, 802.11.

802.11 networks are made up of clients, such as laptops and mobile phones, and infrastructure called **APs** (**access points**) that is installed in buildings. Access points are sometimes called **base stations**. The access points connect to the wired network, and all communication between clients goes through an access point.

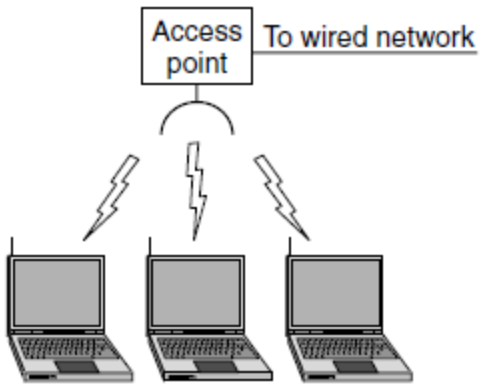

Fig: Wireless network with an access point