Chapter 1 :- Introduction

1.1 DATA COMMUNICATIONS

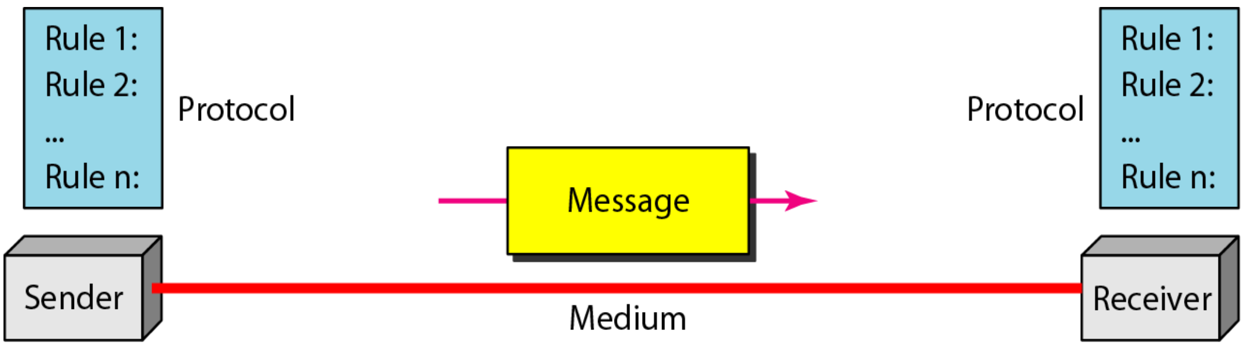
Data communications are the **exchange of data between two devices** via some form of **transmission medium (wired or wireless**). For data communications, the communicating devices must be part of a communication system.

The effectiveness of a data communications system depends on four fundamental characteristics: **delivery, accuracy, timeliness and jitter**.

1. **Delivery:** The system must deliver data to the **correct destination**. Data must be received by the intended device or user.
2. **Accuracy:** The system must deliver the **data accurately**. Data that have been altered in transmission and left uncorrected are unusable.
3. **Timeliness:** The system must deliver data in **a timely manner**. Data delivered late are useless. In the case of video and audio, timely delivery means delivering data as they are produced, in the same order and **without significant delay.**
4. **Jitter:** Jitter refers to the **variation in the packet arrival time**. It is the uneven delay in the delivery of audio or video packets. For example, assume that video packets are sent every 30 ms. If some of the packets arrive with 30-ms delay and others with 40-ms delay, an uneven quality in the video is the result.

Components:

A data communications system has **five** components is as shown below diagram.



1. **Message:** Is the **information (data) to be communicated**. Popular forms of information **include text, numbers, pictures, audio, and video.**
2. **Sender:** Is the device that **sends the data message**. It can be **a computer, workstation, telephone handset, video camera** etc.
3. **Receiver:** Is the device that **receives the message**. It can be a computer, **workstation, telephone handset, television** etc.

4. **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**: 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.**

Data Representation

Information may come in different forms such as **text, numbers, images, audio and video.**

Text: Represented as **bit pattern**, a sequence of bits (0s or 1s). 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**. Most prevalent coding system is called Unicode.
Numbers: Represented as **bit patterns**. However, a code such as ASCII is not used to represent numbers.
Images: Represented as **bit patterns**. In its simplest form, an image is composed of a matrix of **pixels**, where each pixel is a small dot. The size of the pixel depends on the **resolution.**

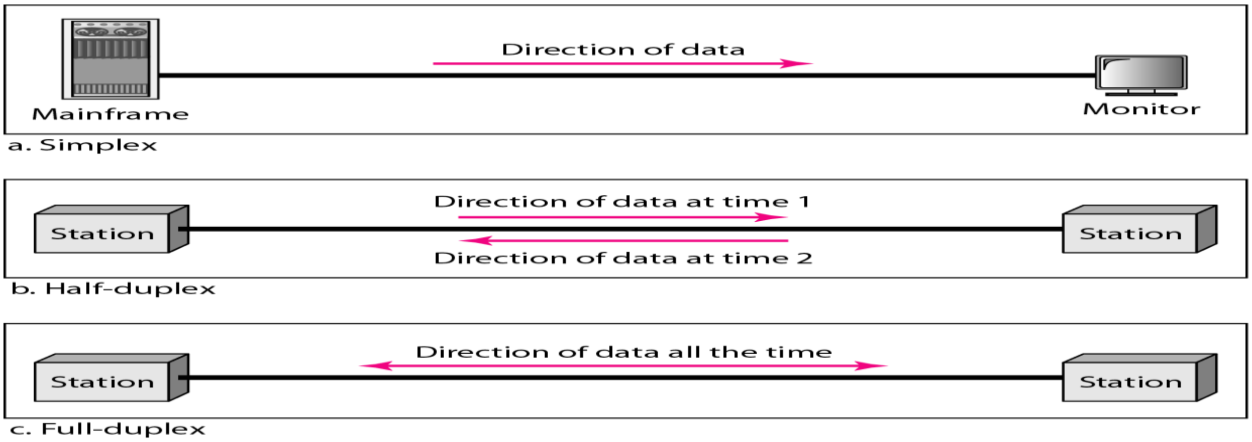
After an image is divided into pixels, each pixel is assigned a bit pattern. The size and the value of the pattern depend on the image. For an image made of only black- and-white dots, a 1-bit pattern is enough to represent a pixel.

There are several methods to represent color images. One method is called RGB(red, green, and blue).
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.
Video: Video refers to the recording or broadcasting of **a picture or movie**. Video can either be produced as a continuous entity, or it can be a combination of images, each a discrete entity, arranged to convey the idea of motion.

Data Flow

Communication between two devices can be simplex, half-duplex, or full-duplex as shown in figure.
Simplex: - The communication **is unidirectional**, as on a one-way street. **Only one of the two devices on a ink can transmit** the other can only **receive**. Keyboards and traditional monitors are examples of simplex devices. The keyboard can only introduce input, the monitor can only accept output.
Half-Duplex: - 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. 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.
Full-Duplex: - Both stations can **transmit and receive simultaneously**. 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. One common example of full-duplex communication is the telephone network.



1.2 NETWORKS

A network is a set of devices (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.

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 process, separate computers handle a subset.

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.

Throughput is the average rate of *successful* message delivery over a communication channel. **The delay** of a network *specifies how long it takes* for a bit of data to travel across the network from one node or endpoint to another.

Reliability

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 are **privacy, Integrity, Authentication and Non repudiation.**

i. **Privacy:** It must not be possible for the credit card number to be **stolen** while on its way to server
ii. **Integrity:** It must not be possible for the credit card number to be **Modified** while on its way to server
iii. **Authentication:** It must be possible for the purchaser and seller to be certain of each other's identity
iv. **Non repudiation**: It must be possible to legally prove that the message was actually sent and received

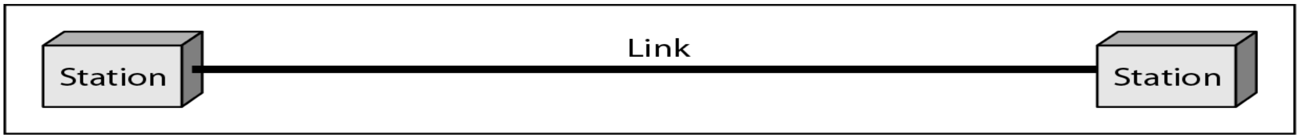
Physical Structures

Type of Connection

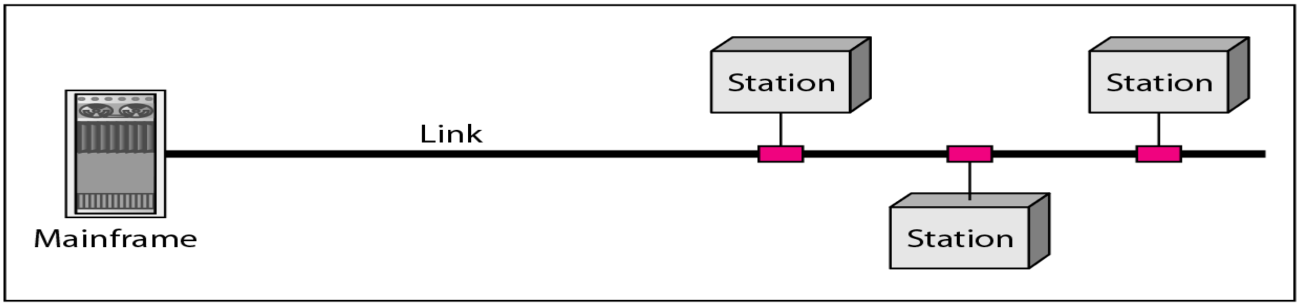
A network is two or more devices connected through links. A link is a communications pathway that transfers data from one device to another. Two devices must be connected in some way to the same link at the same time. There are two possible types of connections: **point-to-point** and **multipoint.**

Point-to-Point:- 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 as shown below figure.

Multipoint:- 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 as shown below figure.



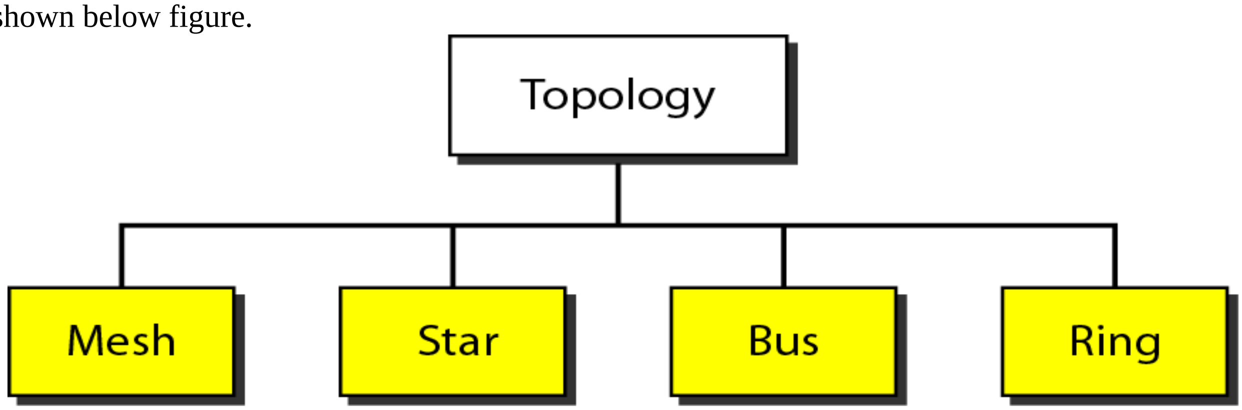
a. Point-to-point



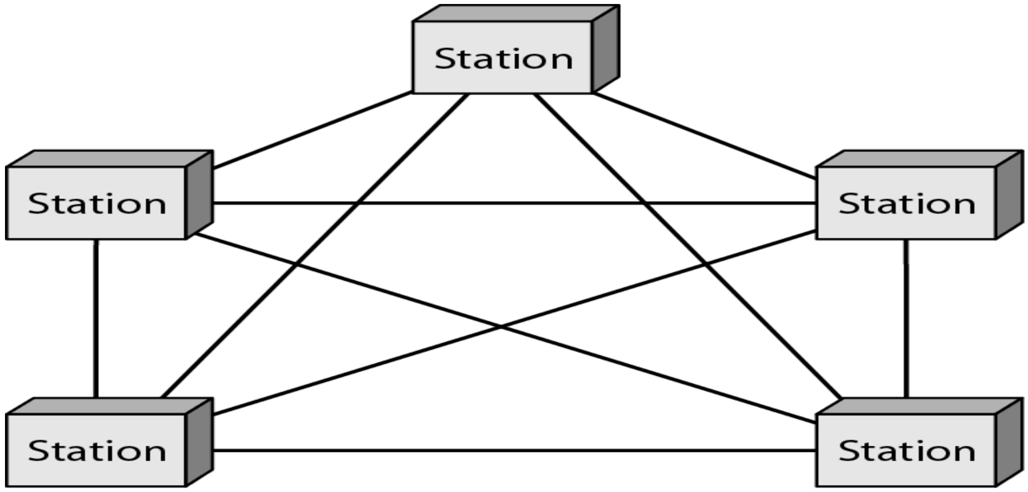
b. Multipoint

Physical Topology

The term physical topology refers to the way in which a network is **laid out physically**. Two 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 (nodes) to one another.** There are four basic topologies possible: ***mesh, star, bus and ring*** as shown below figure.



Mesh: **-E**very 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 say that in a mesh topology, we need **n(n-1)/2 duplex-mode links** as shown below figure.



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. **Is robust**. If one link becomes unusable, it does not incapacitate the entire system.
3. Point-to-point links make fault identification and **fault isolation easy**.

Disadvantages

1. Every device must be connected to all other devices, **installation and reconnection are difficult.**
2. The sheer **bulk of the wiring** can be greater than the available space can accommodate.
3. The hardware required to connect each link can be prohibitively **expensive.**

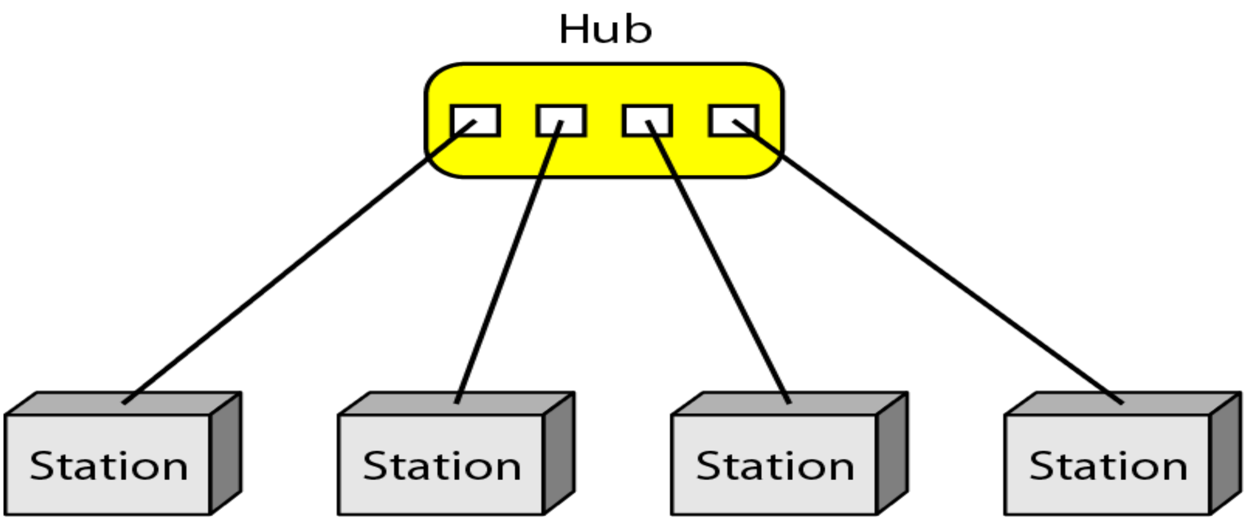
Star Topology: - Each device has a **dedicated point-to-point link only to a central controller, usually called a hub**. The devices are not directly linked to one another. Unlike a mesh topology**, a star topology does not allow direct traffic between devices**. The controller acts as an exchange. If 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 as shown below figure.

Advantages:-

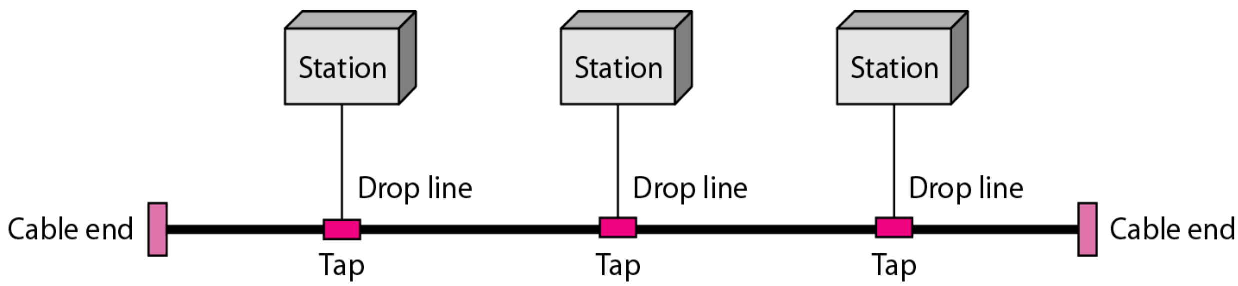
1. A star topology is less expensive than a mesh topology. In a star, each device needs only one link and one I/O port to connect it to any number of others. This factor also makes it easy to install and reconfigure.
2. It is robustness. If one link fails, only that link is affected. All other links remain active. This factor also lends itself to easy fault identification and fault isolation.

Disadvantages:-

1. 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.



Bus Topology: -Is multipoint. One long cable acts as a backbone to link all the devices in a network as shown below figure.



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.

Advantages:

1. **Ease of installation**. Backbone cable can be laid along the most efficient path, and then connected to the nodes by drop lines of various lengths. A bus uses less cabling than mesh or star topologies.
2. **Redundancy is eliminated**. Only the backbone cable stretches through the entire facility.

Disadvantages:

1. 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.
2. Signal reflection at the taps can cause degradation in quality. Adding new devices may therefore require modification or replacement of the backbone.

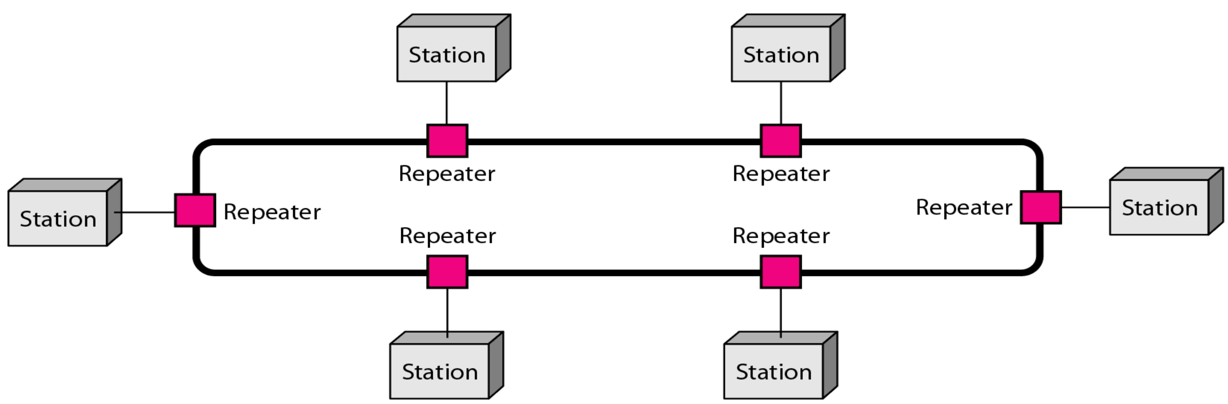
Ring Topology: - Each **device has a dedicated point-to-point connection with only the two devices on either side of it.** A signal is passed along the ring in one direction, from device to device, until it reaches its destination. Each device in the ring incorporates a repeater. When a device receives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

Advantages:

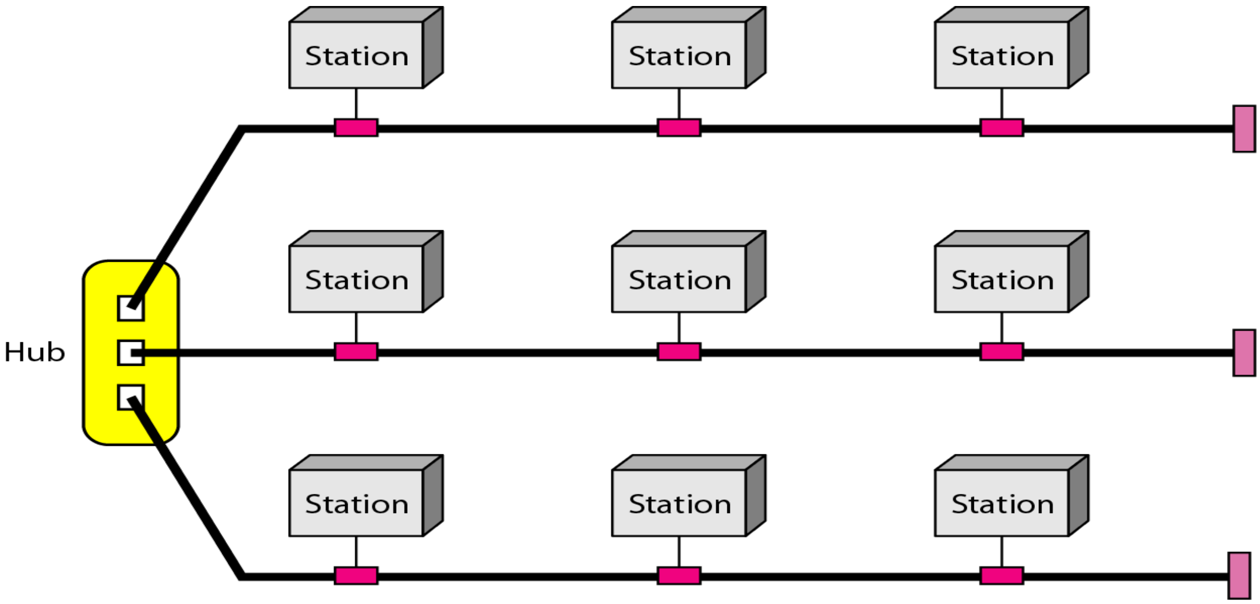
1. A ring is relatively **easy to install and reconfigure**. Each device is linked to only its **immediate neighbors.**
2. To add or delete a device requires changing only two connections.
3. Here **fault isolation is simplified**. Generally in a ring, a signal is circulating at all times. If one device does not receive a signal within a specified period, it can issue an alarm.

Disadvantages:

Unidirectional traffic can be a disadvantage. In a simple ring, a **break in the ring can disable the entire network.**



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:



Network Models

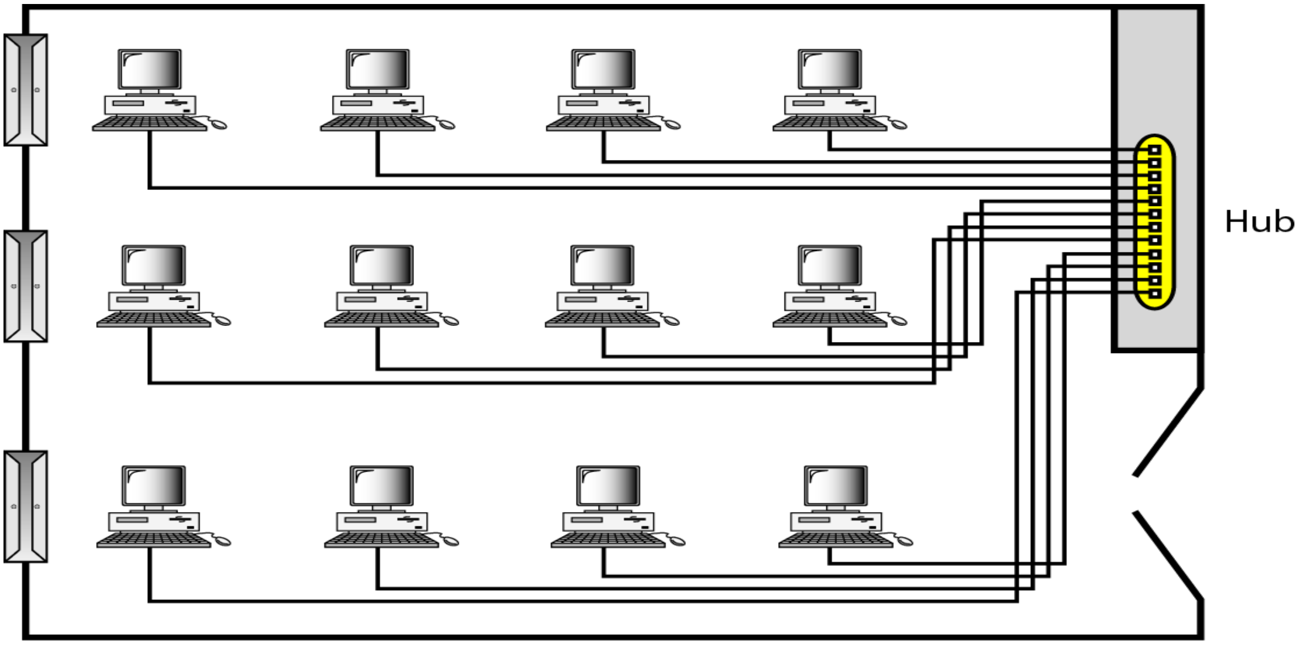
Computer networks are created by different entities. Standards are needed so that these heterogeneous networks can communicate with one another. The two best-known standards are the **OSI model and the Internet model**. The OSI (Open Systems Interconnection) model defines a seven-layer network.

Categories of Networks

Local Area Network

A local area network (LAN) is usually **privately owned and links the devices in a single office,** building, or campus. Depending on the needs of an organization and the type of technology used, a LAN can be as simple as two PCs and a printer in someone's home office Currently, **LAN size is limited to a few kilometers**.

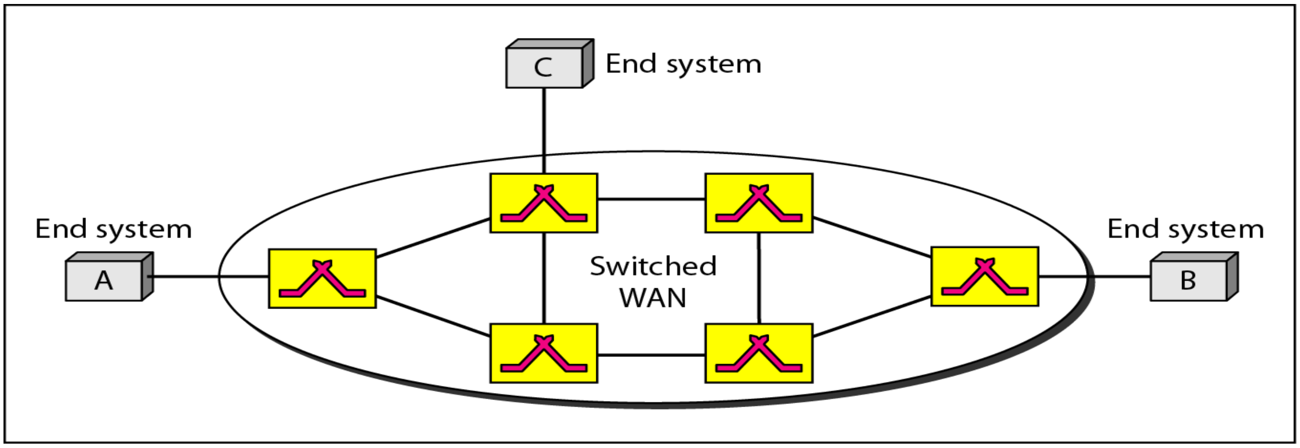
LANs are designed to **allow resources to be shared between personal computers or workstations**. The resources to be shared can include hardware (e.g., a printer), software (e.g., an application program), or data is as shown below figure. A common example of a LAN, found in many business environments, links a workgroup of task-related computers. The most common LAN topologies are bus, ring, and star.



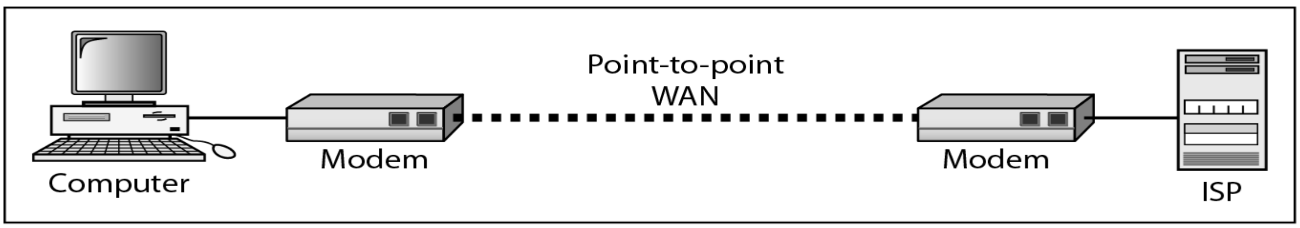
Wide Area Network

A wide area network (WAN) **provides long-distance transmission** of data, image, audio, and video information over large geographic areas that may comprise a country, a continent, or even the whole world.

A WAN can be as **complex as the backbones that connect the Internet** or as simple as a dial-up line that connects a home computer to the Internet. We normally refer to the first as a **switched WAN** and to the second as a **point-to-point WAN**. The switched WAN connects the **end systems, which usually comprise a router** that connects to another LAN or WAN. The point-to-point WAN is normally a **line leased from a telephone or cable TV** provider that connects a home computer or a small LAN to an Internet service provider (ISP). This type of WAN is often used to provide Internet access.



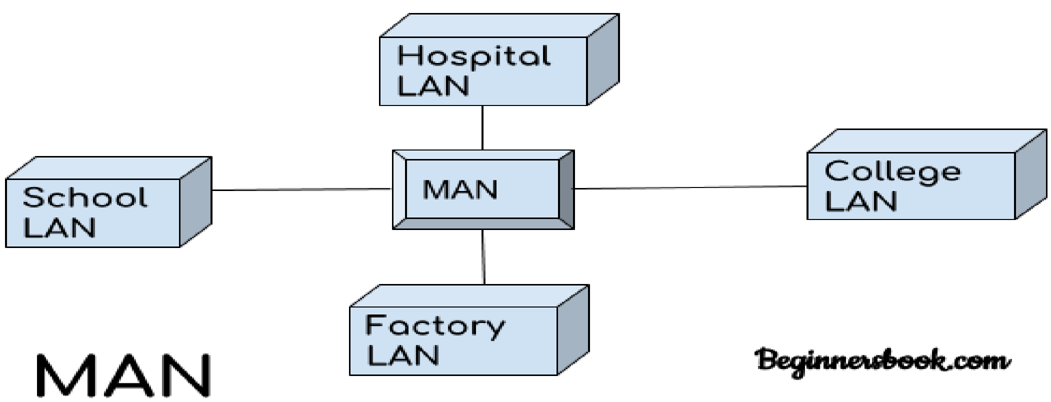
a. Switched WAN



b. Point-to-point WAN

Metropolitan Area Networks

A metropolitan area network (MAN) is a network with a size between a LAN and a WAN. It normally covers the **area inside a town or a city.** It is designed for customers who need high-speed connectivity, normally to the Internet, and have endpoints spread over a city or part of city is as shown below.



Interconnection of Networks:

When two or more networks are connected, they become an internetwork, or internet. As an example, assume that an organization has two offices in separate cities. One established office has **a bus topology** LAN; the other office has **a star topology** LAN. The president lives in some other city and needs to have control over the company from his home. To create a backbone WAN for connecting these three entities a switched WAN has been leased. To connect the LANs to this switched WAN, however, three point-to-point WANs are required. These point-to-point WANs can be a high-speed DSL (Digital Subscriber line) offered by a telephone company or a cable modem line offered by a cable TV provider as shown:

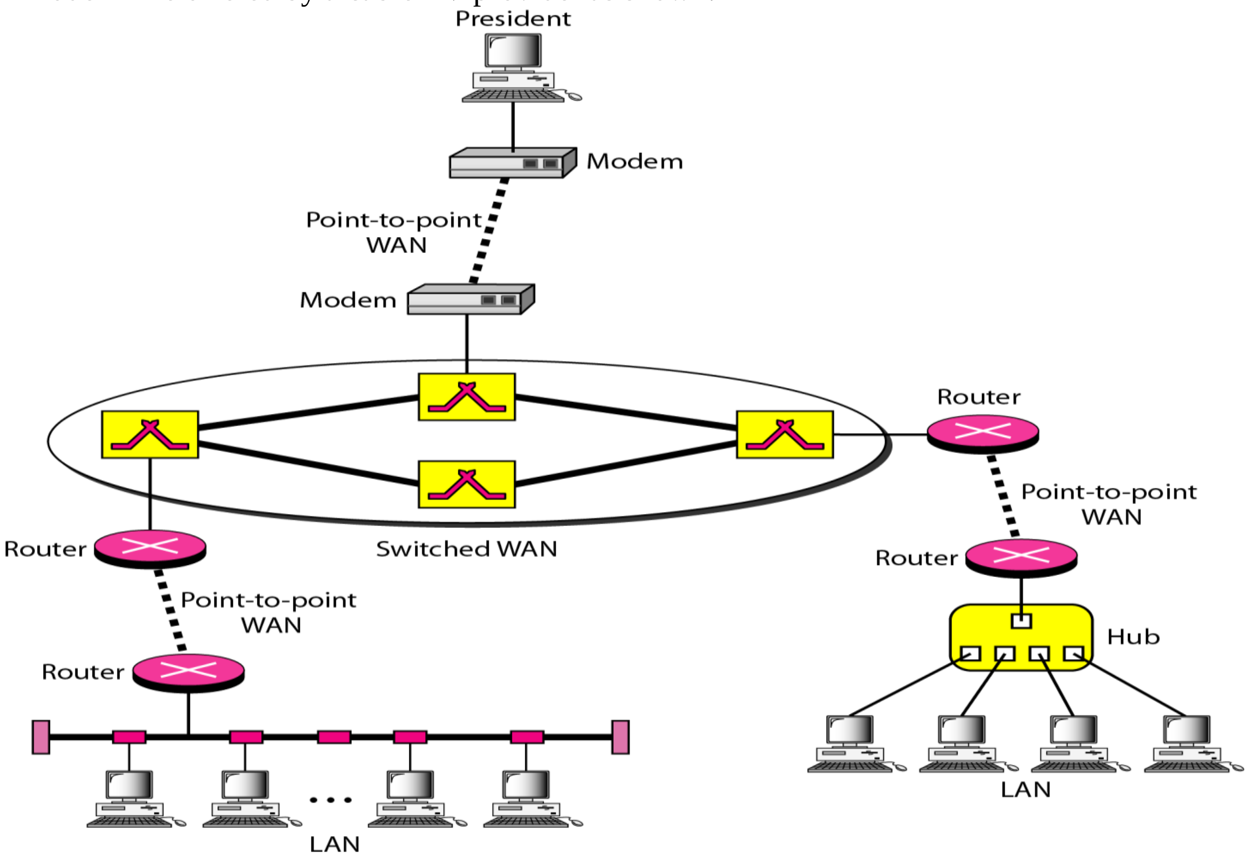


Fig: Heterogeneous network with Four WAN's and Two LAN's

1.3 THE INTERNET

The Internet is a structured, organized system. We begin with a brief history of the Internet.

A Brief History

In the mid-1960s, mainframe computers in research organizations were stand-alone devices. 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). The IMPs, in turn, would be connected to one another.

By 1969, ARPANET was a reality, at the University of California at Los Angeles (UCLA), the University of California at Santa Barbara (UCSB), Stanford Research Institute (SRI), and the University of Utah, 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 Internet Project. Cerf and Kahn's land-mark 1973 paper outlined the protocols to achieve end-to-end delivery of packets.

Shortly thereafter, authorities made a decision to split TCP into two protocols: Transmission Control Protocol (TCP) and Internetworking Protocol (IP). IP would handle datagram routing while TCP would be responsible for higher-level functions such as segmentation, reassembly, and error detection. The internetworking protocol became known as TCP/IP.

The Internet Today

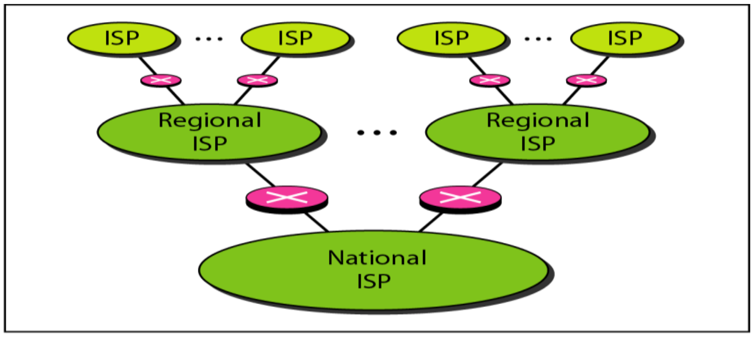
The Internet today is not a simple hierarchical structure. It is made up of many wide- and local-area networks joined by connecting devices and switching stations. It is difficult to give an accurate representation of the Internet because it is continually changing--new networks are being added, existing 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 figure shows a conceptual (not geographic) view of the Internet.

International Internet Service Providers: At the top of the hierarchy are the international service providers that connect nations together.

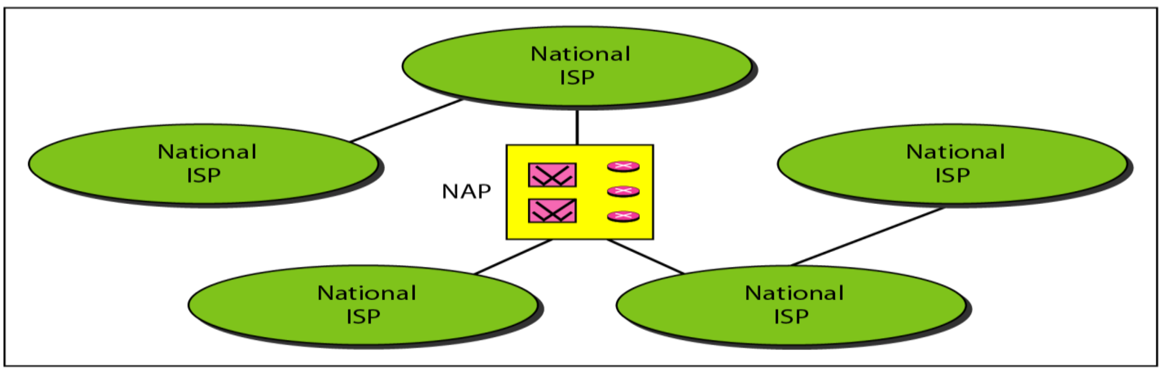
National Internet Service Providers: The national Internet service providers are backbone networks created and maintained by specialized companies. To provide connectivity between the end users, these backbone networks are connected by complex switching stations (normally run by a third party) called 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.

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



a. Structure of a national ISP



b. Interconnection of national ISPs

.

Accessing the Internet

Using Cable Networks

- ✓ A residence can be connected to the Internet by using cable service.
- ✓ Cable service provides a higher speed connection.
- ✓ The speed varies depending on the number of neighbors that use the same cable

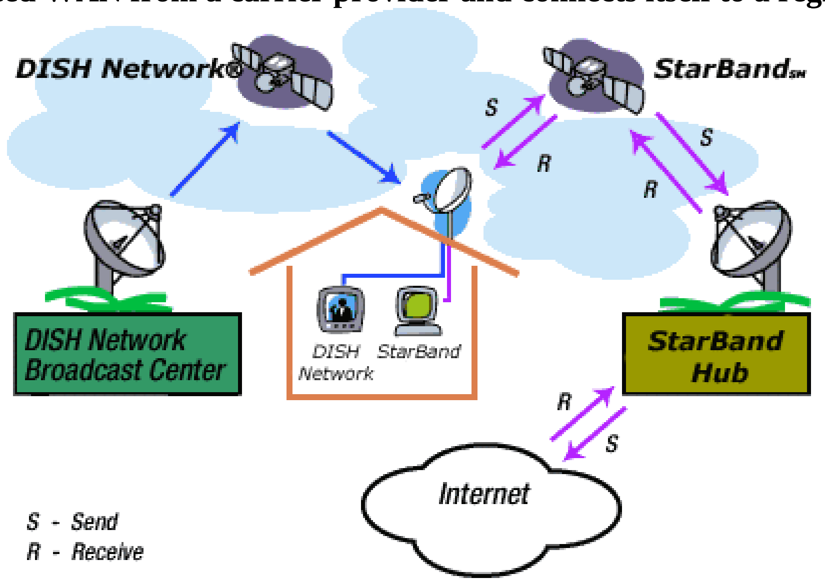
Using Wireless Networks

- ✓ A residence can use a combination of wireless and wired connections to access the Internet.
- ✓ A residence can be connected to the Internet through a wireless WAN.(JioFi WiFi Router Modem)

Direct Connection to the Internet

- ✓ A large organization can itself become a local ISP and be connected to the Internet.
- ✓ The organization

leases a high-speed WAN from a carrier provider and connects itself to a regional ISP.



1.4 PROTOCOLS AND STANDARDS

Protocols

In computer networks, communication occurs between entities in different systems. An entity is anything capable of sending or receiving information. However, two entities cannot simply send bit streams to each other and expect to be understood. For communication to occur, **the entities must agree on a protocol**. **A protocol defines what is communicated, how it s communicated, and when it is communicated**. The key elements of a protocol are **syntax, semantics, and timing**.

• **Syntax**. The term syntax refers to the structure or format of the data, meaning the order in which they are presented.
• **Semantics**. The word semantics refers to the meaning of each section of bits. How is a Particular pattern to be interpreted and specifies what action is to be taken based on that interpretation.
• **Timing**. The term timing refers to two characteristics: when data should be sent and how fast they can be sent.

Standards

Standards are essential in creating and maintaining an open and competitive market for equipment manufacturers and in guaranteeing national and international interoperability of data and telecommunications technology and processes.

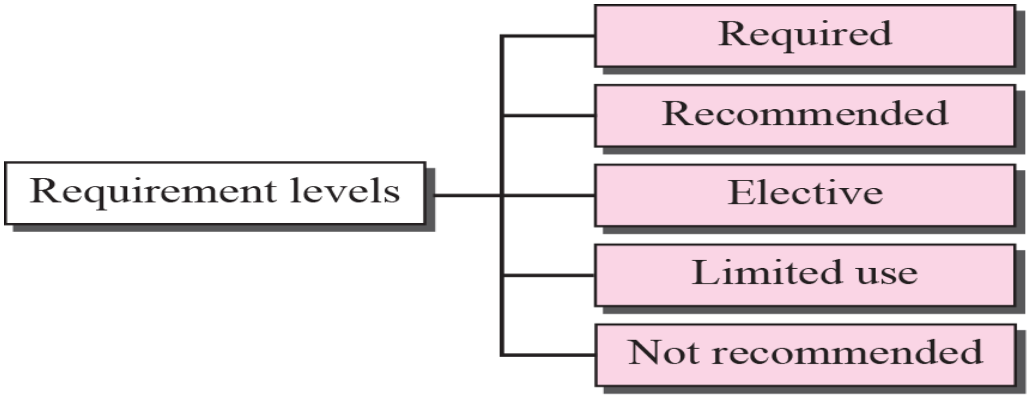
• **De facto**. Standards that have not been approved by an organized body but have been adopted as standards through widespread use are de facto standards.
• **De jure**. Those standards that have been legislated by an officially recognized body are de jure standards.

Standards Organizations
Standards are developed through the cooperation of standards creation committees, forums, and government regulatory agencies.
Standards Creation Committees
While many organizations are dedicated to the establishment of standards, data telecommunications in North America rely primarily on those published by the following:
• **International Organization for Standardization (ISO)**. The ISO is a multinational body whose membership is drawn mainly from the standards creation committees of various Governments throughout the world.
• **International Telecommunication Elnion Telecommunication Standards Sector (ITEl-T)**. By the early 1970s, a number of countries were defining national standards for telecommunications, but there was still little international compatibility
• **American National Standards Institute (ANSI)**. Despite its name, the American National Standards Institute is a completely private, nonprofit corporation.
• **Institute of Electrical and Electronics Engineers (IEEE).** The Institute of Electrical and Electronics Engineers is the largest professional engineering society in the world.
• **Electronic Industries Association (EIA)**. Aligned with ANSI, the Electronic Industries Association is a nonprofit organization devoted to the promotion of electronics manufacturing.

Forums
Telecommunications technology development is moving faster than the ability of standards committees to ratify standards.
Regulatory Agencies
All communications technology is subject to regulation by government agencies such as the Federal Communications Commission (FCC) in the United States. The purpose of these agencies is to protect the public interest by regulating radio, television, and wire/cable communications.

Internet Standards
An Internet standard is a thoroughly tested specification that is useful to and adhered to by those who work with the Internet.

Internet Standards
- ✓ An Internet standard is that is useful to and adhered to by those who work with the Internet.
- ✓ Upon recommendation from the Internet authorities, a draft may be published as a Request for Comment (RFC)
- ✓ RFC becomes an Internet Standard (STD)- RFC 821 (SMTP)

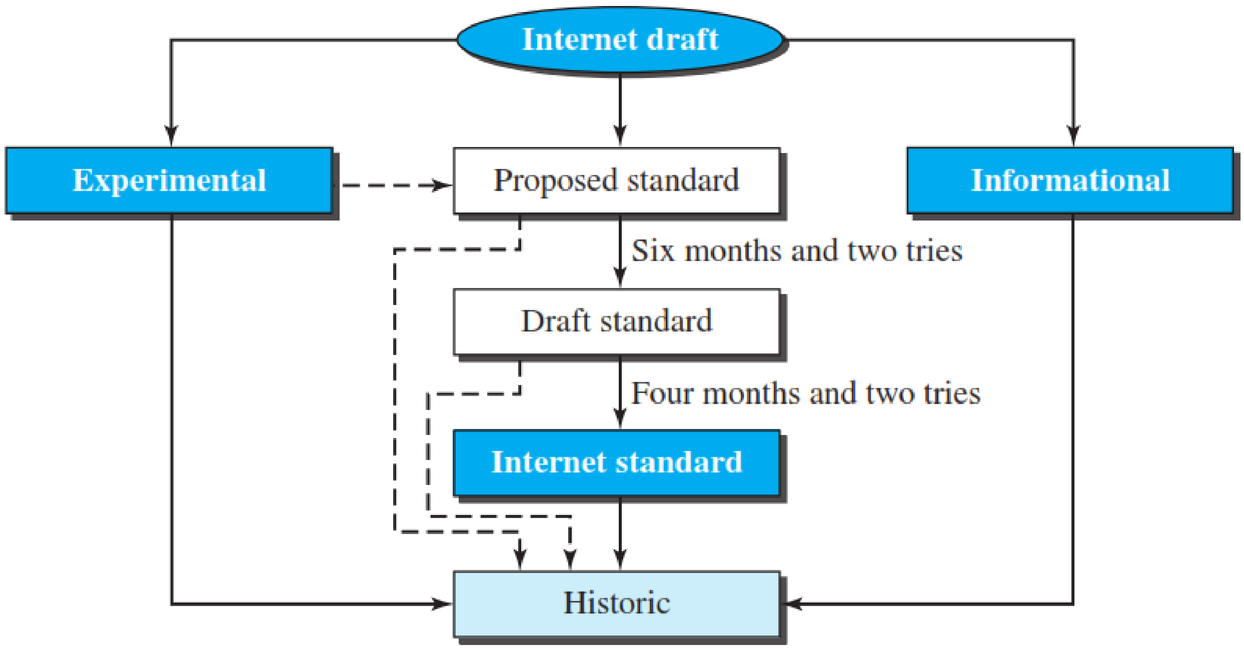


An RFC, six maturity levels: proposed standard, draft standard, Internet standard, historic, experimental, and informational

Maturity Levels

An RFC, during its lifetime, falls into one of six maturity levels: proposed standard, draft standard, Internet standard, historic, experimental, and informational Proposed Standard.

- ✓ A proposed standard is a specification that is stable, well understood, and of sufficient interest to the Internet community. At this level, the specification is usually tested and implemented by several different groups
- ✓ Draft Standard. A proposed standard is elevated to draft standard status after at least two successful independent and interoperable implementations. Barring difficulties, a draft standard, with modifications if specific problems are encountered, normally becomes an Internet standard
- ✓ Internet Standard. A draft standard reaches Internet standard status after demonstrations of successful implementation.
- ✓ Historic. The historic RFCs are significant from a historical perspective. They either have been superseded by later specifications or have never passed the necessary maturity levels to become an Internet standard.
- ✓ Experimental. An RFC classified as experimental describes work related to an experimental situation that does not affect the operation of the Internet. Such an RFC should not be implemented in any functional Internet service.
- ✓ Informational. An RFC classified as informational contains general, historical, or tutorial information related to the Internet. It is usually written by someone in a non-Internet organization, such as a vendor



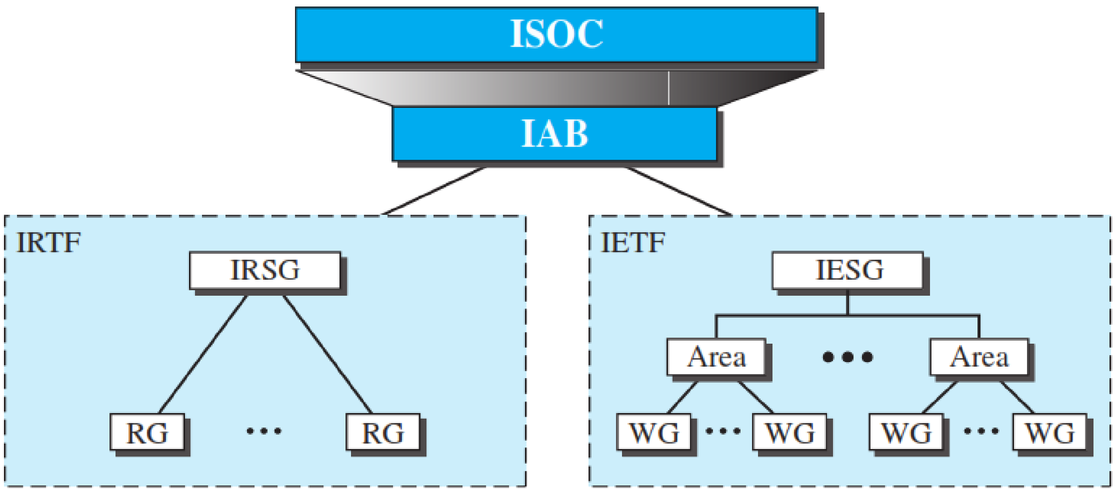
Requirement Levels

RFCs are classified into five requirement levels: required, recommended, elective, limited use, and not recommended.

- ✓ Required. An RFC is labeled required if it must be implemented by all Internet systems to achieve minimum conformance. For example, IP and ICMP (Chapter 19) are required protocols.
- ✓ Recommended. An RFC labeled recommended is not required for minimum conformance; it is recommended because of its usefulness. For example, FTP (Chapter 26) and TELNET (Chapter 26) are recommended protocols.
- ✓ Elective. An RFC labeled elective is not required and not recommended. However, a system can use it for its own benefit.
- ✓ Limited Use. An RFC labeled limited use should be used only in limited situations. Most of the experimental RFCs fall under this category.
- ✓ Not Recommended. An RFC labeled not recommended is inappropriate for general use. Normally a historic (deprecated) RFC may fall under this category.

Internet Administration

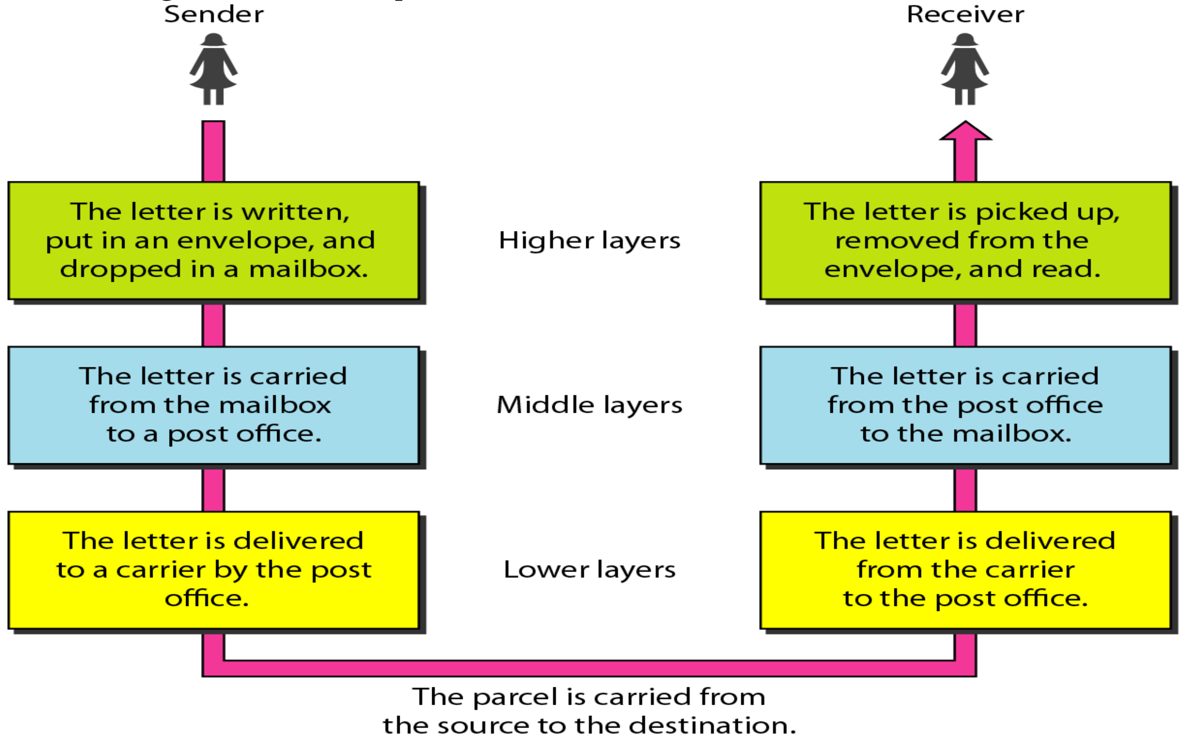
The Internet Society (ISOC) is an international, nonprofit organization formed in 1992 to provide support for the Internet standards process



Chapter 2:- Network Models

2.1 LAYERED TASKS

As an example, let's 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. The figure shows the steps in this task.



Sender, Receiver, and Carrier

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

At the Sender Site

The activities that take place at the sender site, in order, are:

- **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.
- **Middle layer**: - The letter is picked up by a letter carrier and delivered to the post office.
- **Lower layer**: - The letter is sorted at the post office; a carder transports the letter.

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

At the Receiver Site

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

Hierarchy

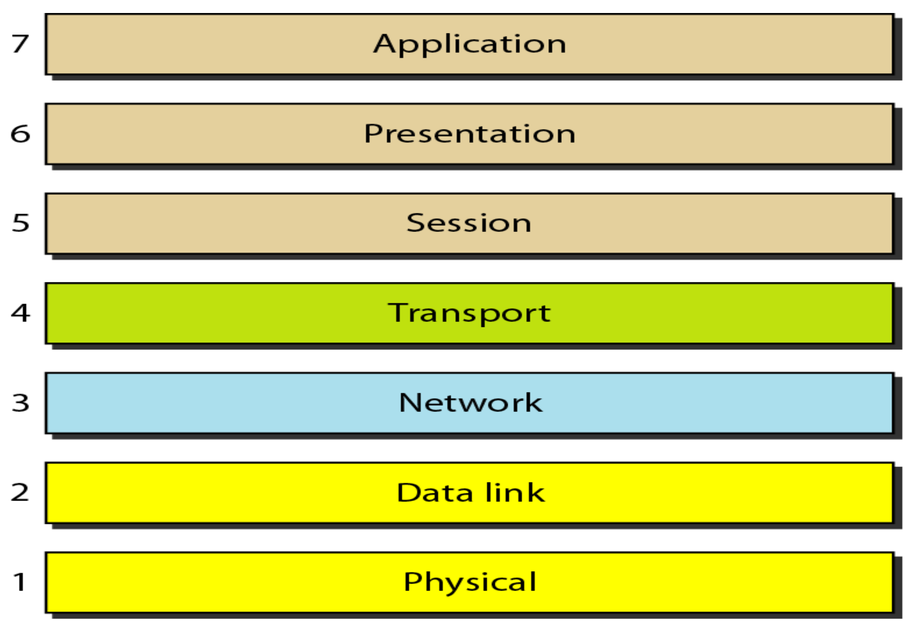
There are three different activities at the sender and another three activities at the receiver. The task of transporting the letter between the sender and the receiver is done **by the carrier**. Something that is not obvious immediately is that the tasks must be done in the order given in the hierarchy. At the sender side, the letter must be written and dropped in the mailbox before being picked up by the letter carrier and delivered to the post office. At the receiver side, the letter must be dropped in the recipient mailbox before being picked up and read by the recipient.

Services

Each layer at the sending site uses **the services of the layer immediately below it.** The sender at the higher layer uses the services of the middle layer. **The lower layer uses the services of the carrier.**

2.2 THE OSI MODEL

The OSI model is a layered framework for the design of network systems that allows communication between **all types of computer systems**. It consists of **seven separate** but related layers, each of which defines a part of the process of moving information across a network is as shown below figure.

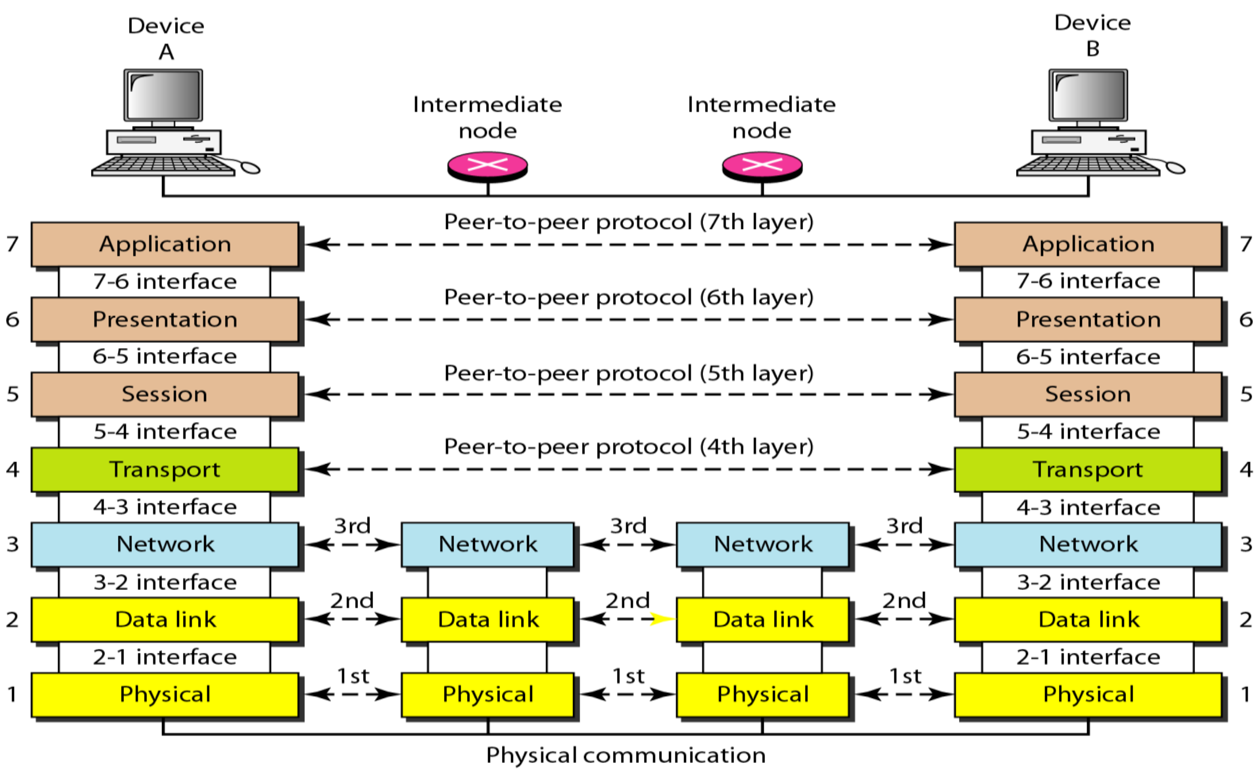


Layered Architecture

The OSI model is composed of seven ordered layers: physical (layer 1), data link (layer 2), network (layer 3), transport (layer 4), session (layer 5), presentation (layer 6), and application (layer 7). The below figure shows the layers involved when a m essage is sent from device A to device B. As the message travels from A to B, it may pass through many **intermediate nodes**. These intermediate nodes **usually involve only the first three layers of the OSI model.** Within a single machine, each layer calls upon the services of the layer just below it.

Peer-to-Peer Processes

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



At layer 1 the entire package is converted to some form that can be transmitted to the receiving device. At the receiving machine, the message is unwrapped layer by layer, with each process receiving and removing the data meant for it. For example, layer 2 removes the data meant for it, then passes the rest to layer 3. Layer 3 then removes the data meant for it and passes the rest to layer 4, and so on.

2.3 LAYERS IN THE OSI MODEL

Physical Layer

The physical layer is also concerned with the following:

- **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.**
- **Representation of bits:** The physical layer data **consists of a stream of bits** sequence of 0s or ls) with no interpretation. To be transmitted, bits must be encoded into signals electrical or optical. The physical layer defines the type of encoding.
- **Data rate**: The transmission rate the **number of bits sent each second**. In other words, the physical layer **defines the duration of a bit.**
- **Synchronization of bits**: **The sender and receiver not only must use the same bit rate** but also must be synchronized at the bit level.
- **Line configuration:** The physical layer is concerned with **the connection of devices to the media.** In a point-to-point configuration, two devices are connected through a **dedicated link**. In a multipoint configuration, **a link is shared among several devices.**
- **Physical topology:** The physical topology defines **how devices are connected to make a network**. Devices can be connected by using a mesh topology a star topology a ring topology, a bus topology or a hybrid topology.
- **Transmission mode**: The physical layer also defines the **direction of transmission between two devices: simplex, half-duplex, or full-duplex.**

Data Link Layer

Other responsibilities of the data link layer include the following:

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

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

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

Network Layer

- **Logical addressing:**. If a packet passes the network boundary, we need another addressing system to help distinguish the source and destination systems. The network layer adds a **header to the packet** coming from the upper layer that, among other things, includes the **logical addresses** of the sender and receiver.
- **Routing**: When independent networks 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**. The below figure illustrates end-to-end delivery by the network layer.

Transport Layer

- **Service-point addressing**: Computers often run several programs at the **same time**. For this reason, source-to-destination delivery means delivery not only from one computer to the next but also from a specific process on one computer to a specific process on the other. The transport layer header must therefore **include 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 and to identify and replace packets that were lost in transmission.
- **Connection control**: The transport layer can be either **connectionless or connection oriented**. A connectionless transport layer treats each segment as an independent packet and delivers it to the transport layer at the destination machine. A connection-oriented transport layer makes a connection with the transport layer at the destination machine first before delivering the packets. After all the data are transferred, the connection is terminated.
- **Flow control:** flow control at this layer is performed end to end rather than across a single link.
- **Error control: I**s performed **process-to-process** rather than across a single link. The sending transport layer **makes sure that the entire message arrives at the receiving transport layer without error**. Error correction is usually achieved through retransmission.

Session Layer

- **Dialog control**: It allows the **communication between two processes to take place in either half- duplex or full-duplex 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.

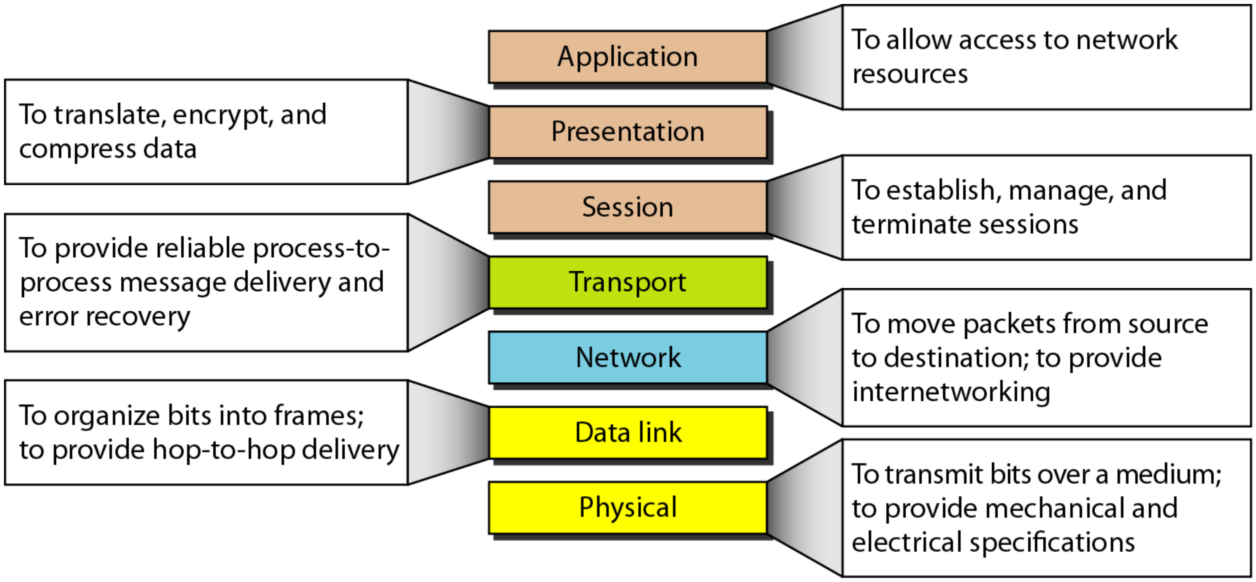
Presentation Layer

Specific responsibilities of the presentation layer include:

- **Translation:** The processes in two systems are usually **exchanging information in the form of character strings, numbers, and so on**. The information must be **changed to bit streams before being transmitted**. Because different computers use different encoding systems
- **Encryption:** Encryption means that the sender transforms the **original information** to **another form** and **sends the resulting message out over the network**. Decryption reverses the original process to transform the message back to its original form.
- **Compression:** Data compression **reduces the number of bits contained** in the information. Data compression becomes particularly important in the transmission of multimedia **such as text, audio, and video**.

Application Layer

- **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 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.



2.4 TCP/IP PROTOCOL SUITE

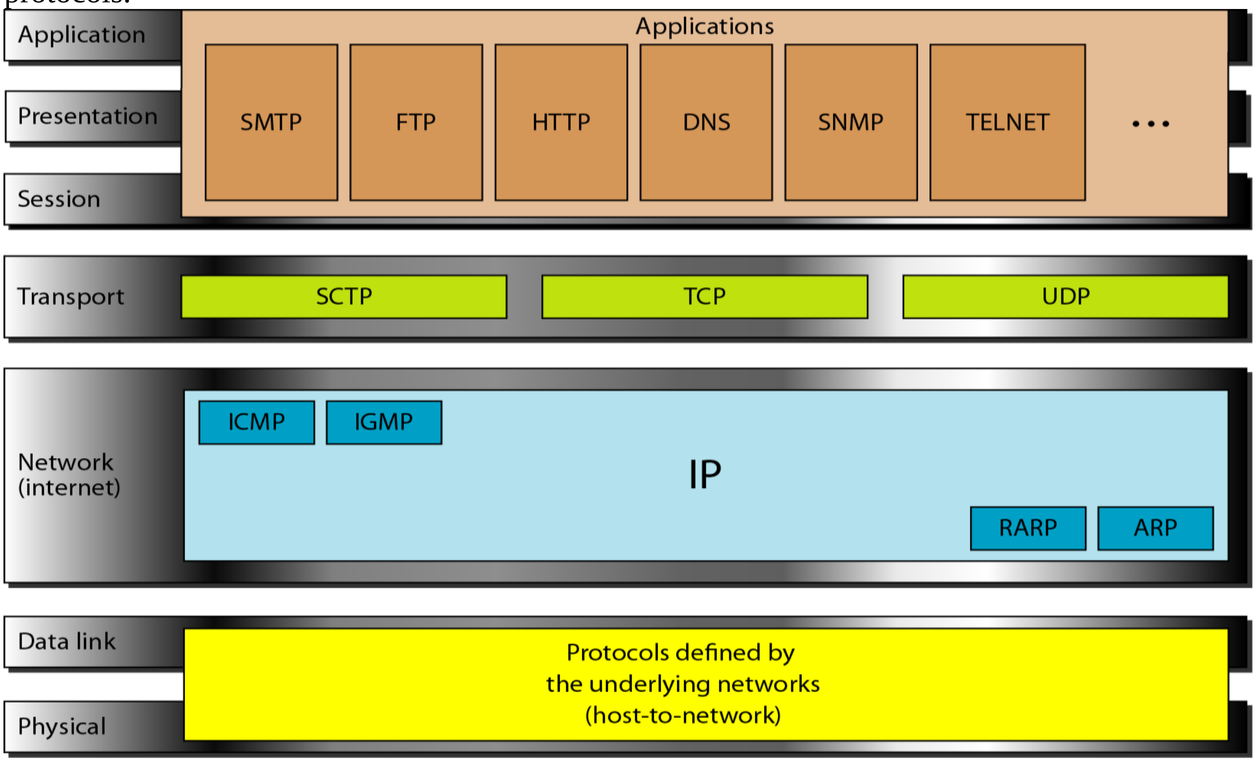
The TCP/IP protocol suite defined in **four layers*: host-to-networks, internet, transport and application.*** Here host to network is equivalent to **combination of physical and data link layer.** The first **four layers provide physical standards, network interfaces, internetworking, and transport functions** that correspond to the first four layers of the OSI model. The three topmost layers in the OSI model, however, are represented in TCP/IP by a single layer called the application layer.

TCP/IP is a hierarchical protocol made up of interactive modules, each of which provides a specific functionality; however, the modules are not necessarily interdependent. Whereas the OSI model specifies which functions belong to each of its layers, **the layers of the TCP/IP protocol suite contain relatively independent protocols that can be mixed and matched depending on the needs of the system**.

At the transport layer, TCP/IP defines three protocols: **Transmission Control Protocol (TCP), User Datagram Protocol (UDP), and Stream Control Transmission Protocol (SCTP).** At the network layer, the main protocol defined by **TCP/IP is the Internetworking Protocol (IP.**

Physical and Data Link Layers

TCP/IP does not define any **specific protocol.** It supports all the standard and proprietary protocols.



Network Layer

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

Internetworking Protocol (IP)

The Internetworking Protocol (IP) is the transmission mechanism used by the TCP/IP protocols. **It is an unreliable and connectionless protocol a best-effort delivery service. IP assumes the unreliability of the underlying layers and does its best to get a transmission through to its destination, but with no guarantees.** IP transports data in packets called datagrams, each of which is transported separately. Datagrams can travel along different routes and can arrive out of sequence or be duplicated. IP does not keep track of the routes and has no facility for reordering datagrams once they arrive at their destination.

Address Resolution Protocol

The Address Resolution Protocol (ARP) is used **to associate a logical address with a physical address**. 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**.

Internet Control Message Protocol

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

Internet Group Message Protocol

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

Transport Layer

Traditionally the transport layer was represented in TCP/IP by two protocols: ***TCP and UDP. IP is a host-to-host protocol, meaning that it can deliver a packet from one physical device to another. UDP and TCP are transport level protocols responsible for delivery of a message from a process to another process..***

User Datagram Protocol

It is a process-to-process protocol that adds only port addresses, checksum error control, and length information to the data from the upper layer.

Transmission Control Protocol

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

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

Stream Control Transmission Protocol

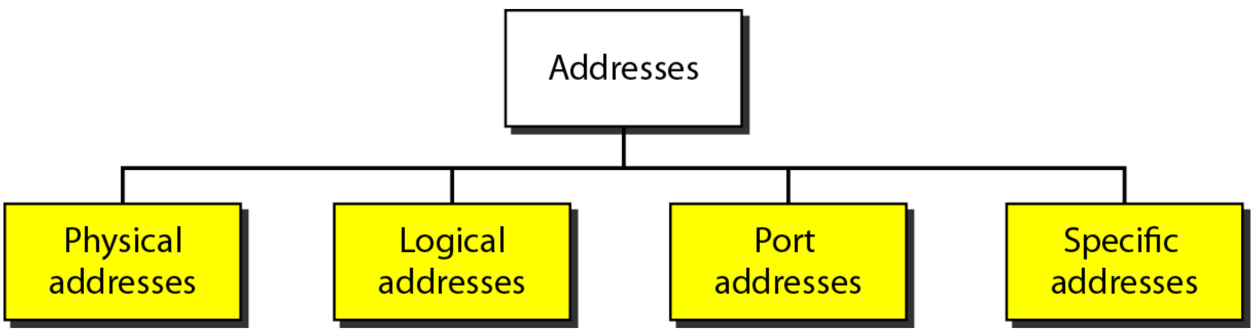
SCTP provides support ***for newer applications such as voice over the Internet***

Application Layer

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

ADDRESSING

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

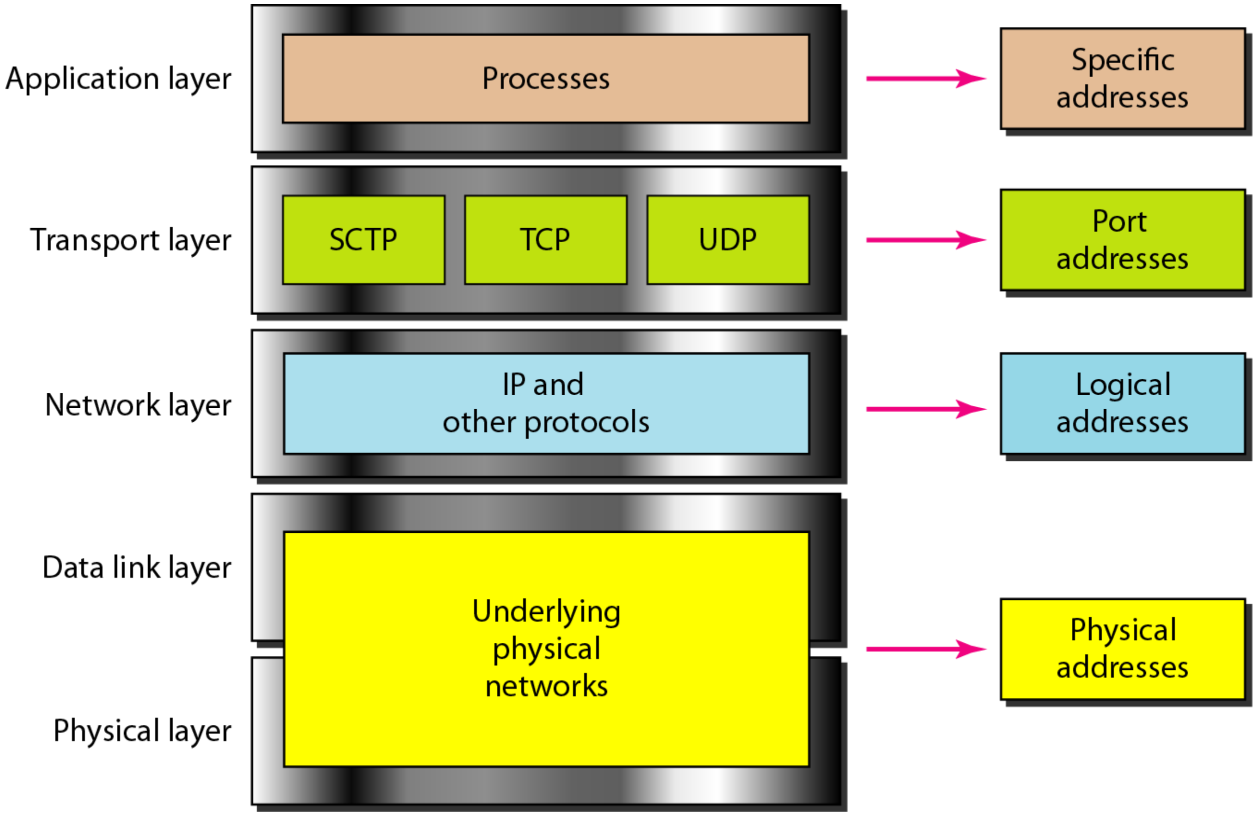


Physical Addresses

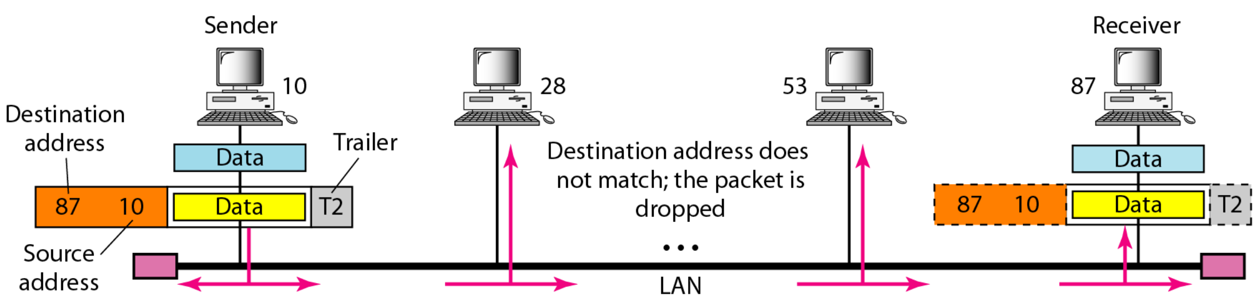
The physical address, also known as the link address, is the address of a node as defined by its LAN or WAN.

It is included in the frame used by the data link layer. It is the lowest-level address. The size and format of these addresses vary depending on the network.

For example, Ethernet uses a 6-byte (48-bit) physical address that is imprinted on the network interface card (NIC).



Below Figure a node with physical address 10 sends a frame to a node with physical address 87. The two nodes are connected by a link (bus topology LAN). As the figure shows, the computer with physical address 10 is the sender, and the computer with physical address 87 is the receiver.

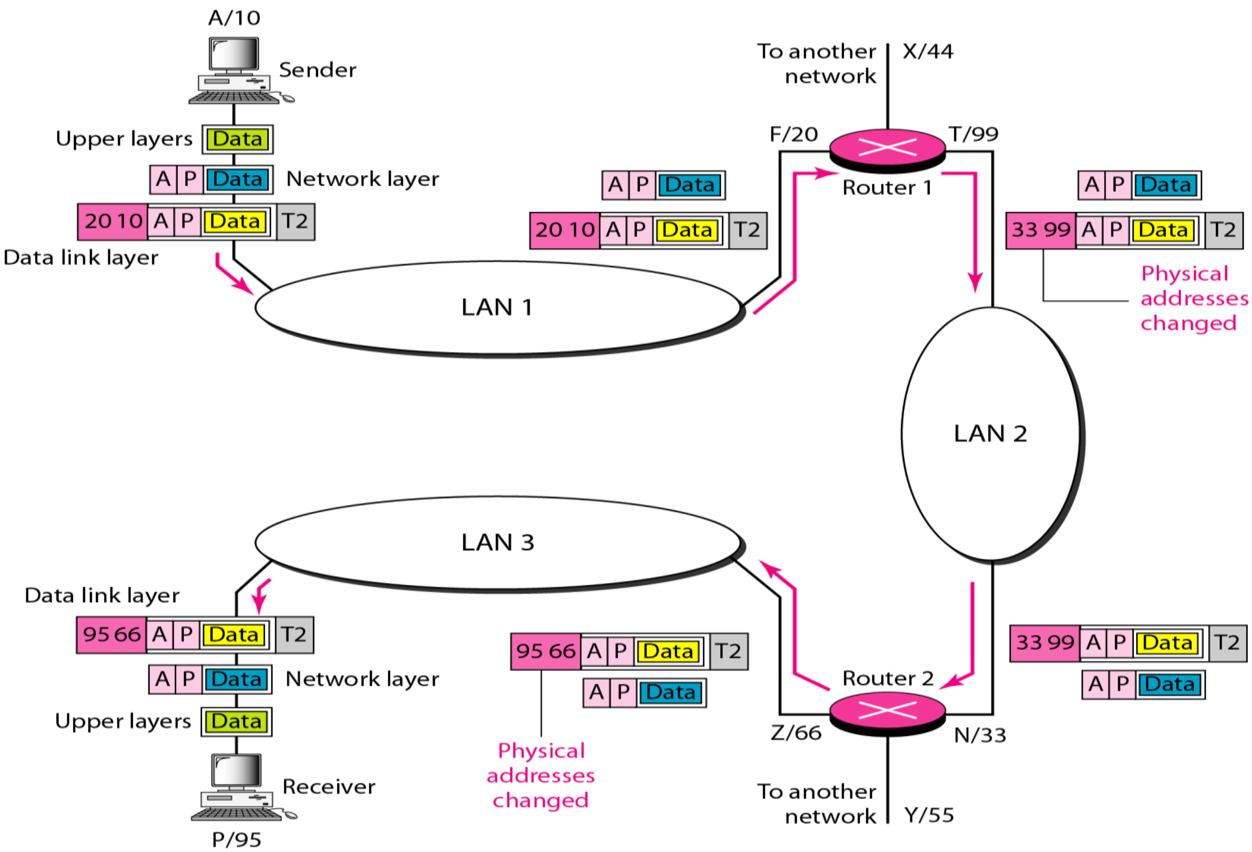


Most local-area networks use a 48-bit (6-byte) physical address written as 12 hexadecimal digits; every byte (2 hexadecimal digits) is separated by a colon, as: 07:01:02:01:2C:4B

Logical Addresses

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

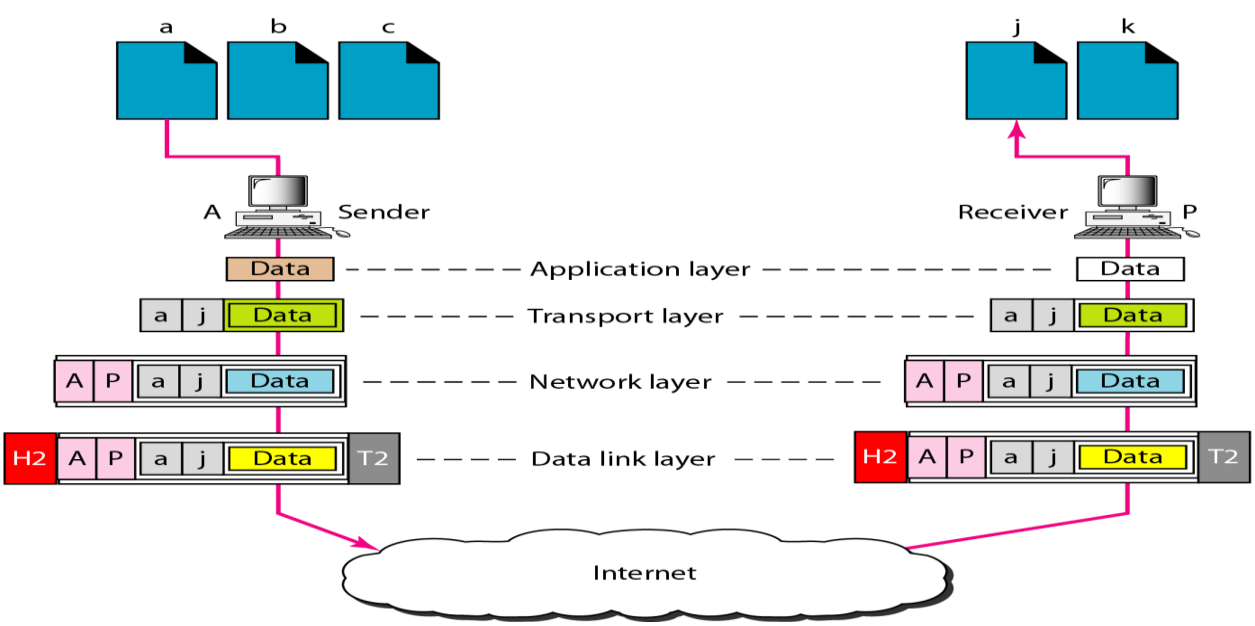
Below Figure shows a part of an internet with two routers connecting three LANs. Each device (computer or router) has a pair of addresses (logical and physical) for each connection. In this case, each computer is connected to only one link and therefore has only one pair of addresses. Each router, however, is connected to three networks (only two are shown in the figure). So each router has three pairs of addresses, one for each connection.



Port Addresses

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

Below Figure shows two computers communicating via the Internet. The sending computer is running three processes at this time with port addresses a, b, and c. The receiving computer is running two processes at this time with port addresses j and k. Process a in the sending computer needs to communicate with process j in the receiving computer. Note that although physical addresses change from hop to hop, logical and port addresses remain the same from the source to destination.



Chapter 3 :- Data and Signals

ANALOG AND DIGITAL

Data must be transformed to electromagnetic signals. Data can be analog or digital. Analog data are **continuous** and take continuous values. Digital data have **discrete states** and take discrete values. Signals can also be analog or digital. Analog signals can have an **infinite number of values in a range, digital signals can have only a limited number of values.**

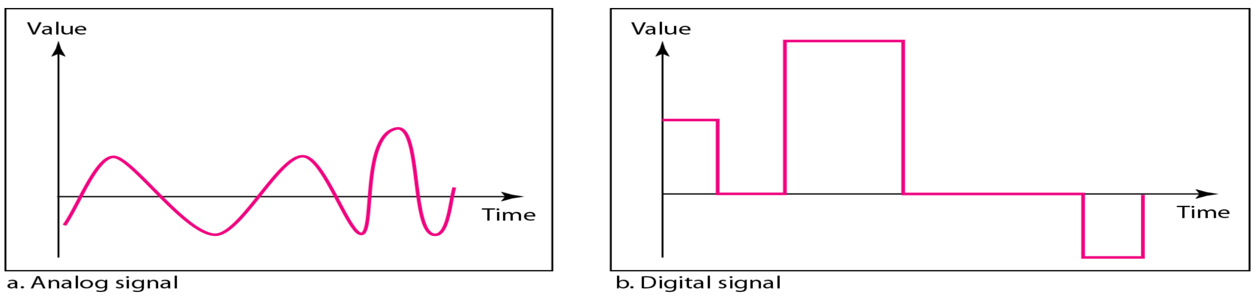


Figure 3.1Analog and Digital Signal

Both analog and digital signal can be **periodic or non periodic**. A periodic signal **completes a pattern within a measurable time frame called period and repeats the pattern over subsequent identical period. Completion a full pattern is called a cycle. A signal which does not repeat itself after a specific interval of time is called non periodic.**

PERIODIC ANALOG SIGNALS

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

SINE WAVE

The sine wave is the most fundamental form of a periodic analog signal. A sine wave can be represented by three parameters: **peak amplitude, frequency and phase.** Below figure shows the sine wave, each cycle consists of single arc below and above the time axis.

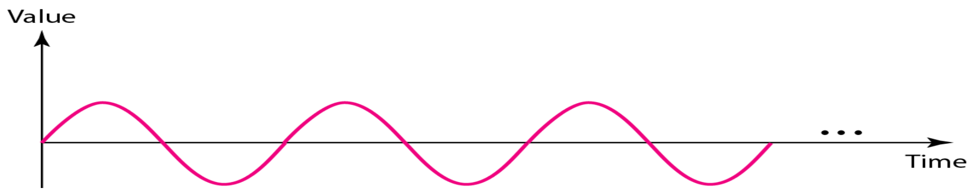


Figure 3.2 ***A sine wave***

PEAK AMPLITUDE:

Peak amplitude is the absolute value of highest intensity, proportional to energy it carries. Peak is the maximum value, either positive or negative, that a waveform attains. Peak values can be expressed for voltage, current, or power is as shown below.

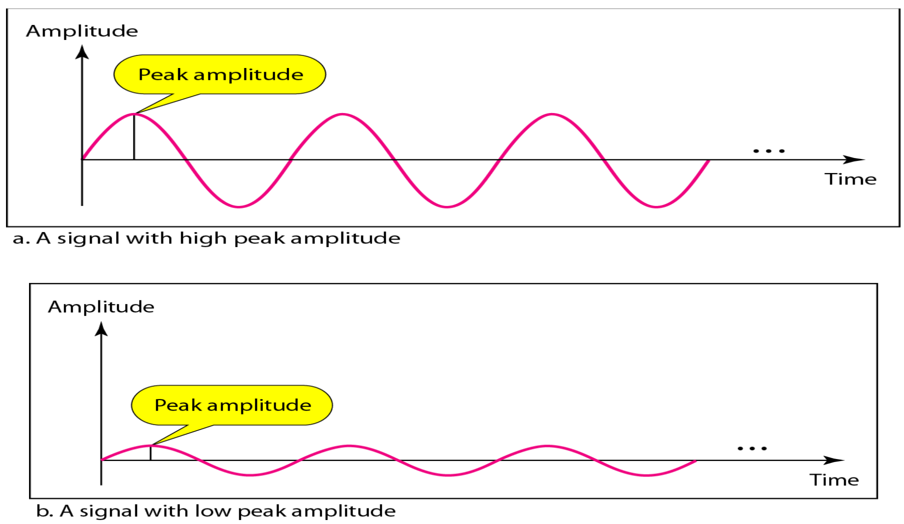


Figure 3.3 Two *signals with the same phase and frequency, but different amplitudes*

PERIOD AND FREQUENCY:

Frequency and period are the inverse of each other. **Frequency is the rate of change with respect to time.** Change in a **short span of time means high frequency**. Change over a **long span of time means low frequency** is as shown below figure 3.4. The frequency and period are inversely proportional ie.

$$f = \frac{1}{T} \quad \text{and} \quad T = \frac{1}{f}$$

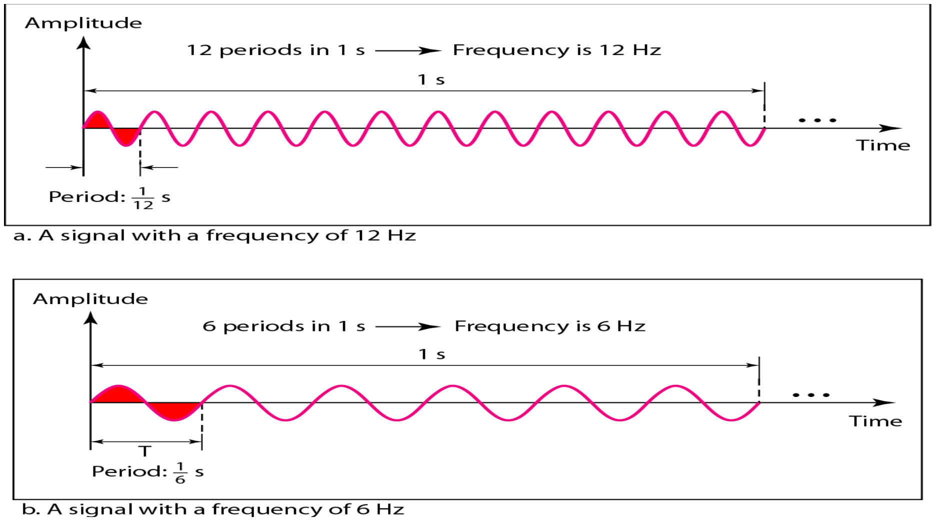


Figure 3.4 Two *signals with the same amplitude and phase, but different frequencies*

Table 3.1 Units *of period and frequency*

Unit	*Equivalent*	*Unit*	*Equivalent*
Seconds (s)	1 s	Hertz (Hz)	1 Hz
Milliseconds (ms)	10^{-3} s	Kilohertz (kHz)	10^3 Hz
Microseconds (μs)	10^{-6} s	Megahertz (MHz)	10^6 Hz
Nanoseconds (ns)	10^{-9} s	Gigahertz (GHz)	10^9 Hz
Picoseconds (ps)	10^{-12} s	Terahertz (THz)	10^{12} Hz

Example 3.3

The power we use at home has a frequency of 60 Hz. Find period?

$$T = \frac{1}{f} = \frac{1}{60} = 0.0166 \text{ s} = 0.0166 \times 10^3 \text{ ms} = 16.6 \text{ ms}$$

Example 3.4

Express a period of 100 ms in microseconds.

Solution

From Table 3.1 we find the equivalents of 1 ms (1 ms is 10^{-3} s) and 1 s (1 s is 10^6 μs). We make the following substitutions:.

$$100 \text{ ms} = 100 \times 10^{-3} \text{ s} = 100 \times 10^{-3} \times 10^6 \text{ μs} = 10^2 \times 10^{-3} \times 10^6 \text{ μs} = 10^5 \text{ μs}$$

Example 3.5

The period of a signal is 100 ms. what is its frequency in kilohertz?

Solution

First we change 100 ms to seconds, and then we calculate the frequency from the period (1 Hz = 10^{-3} kHz).

$$100 \text{ ms} = 100 \times 10^{-3} \text{ s} = 10^{-1} \text{ s}$$

$$f = \frac{1}{T} = \frac{1}{10^{-1}} \text{ Hz} = 10 \text{ Hz} = 10 \times 10^{-3} \text{ kHz} = 10^{-2} \text{ kHz}$$

FREQUENCY AND PHASE

If a signal does not change at all, its frequency **is zero.** If a signal changes **instantaneously, its frequency is infinite.** Phase describes the position of the waveform relative to time 0.The phase measured in degrees or radians is as shown below figure 3.5.

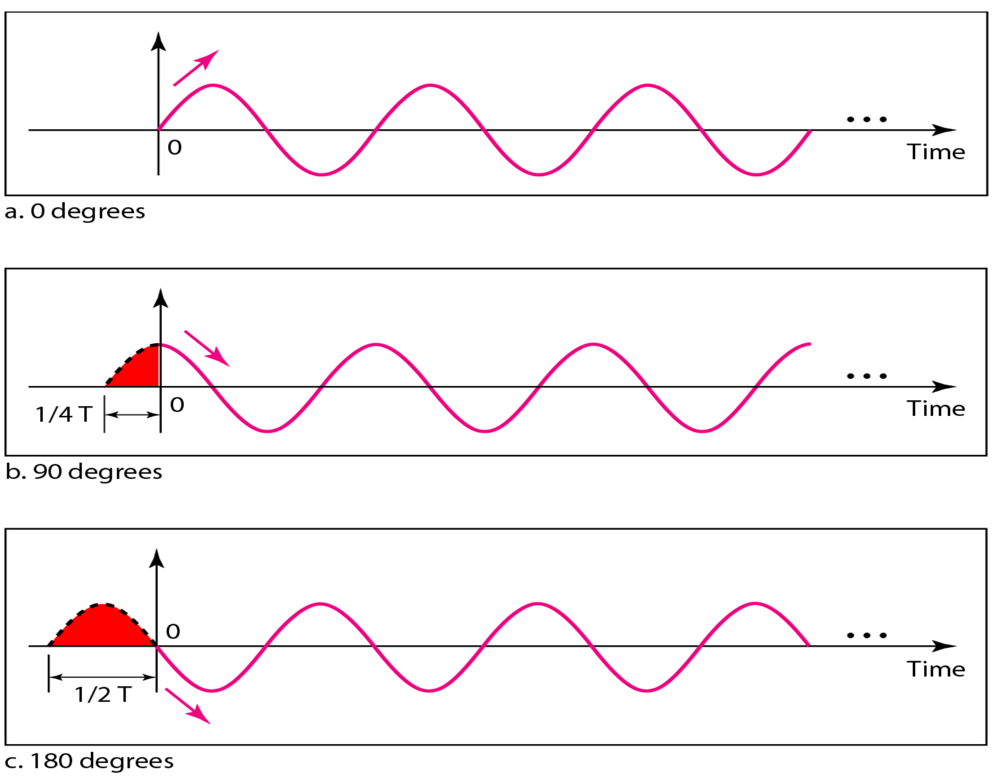


Figure 3.5 ***Three sine waves with the same amplitude and frequency, but different phases***

From above figure

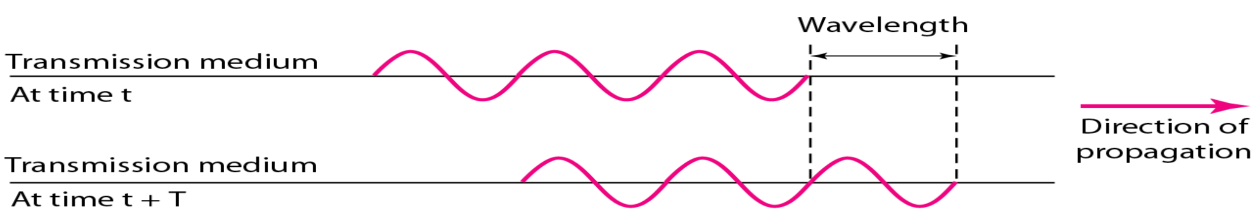
- ***A sine wave with phase of 0 degree stats at time 0 with zero amplitude, and amplitude increasing also not shifted.***
- ***A sine wave with phase of 90 degree stats at time 0 with peak amplitude, and amplitude decreasing also shifted to the left by ¼ cycle.***
- ***A sine wave with phase of180 degree stats at time 0 with zero amplitude, and amplitude decreasing also shifted to the left by ½ cycle.***

WAVELENGTH AND PROPAGATION SPEED

Wavelength: The **distance a simple signal can travel in one period**.

Propagation speed: The rate at which a signal or bit travels; measured by distance/second. **Propagation time**: the time required for a signal to travel from one point to another.

Wavelength = propagation speed x period = propagation speed /frequency



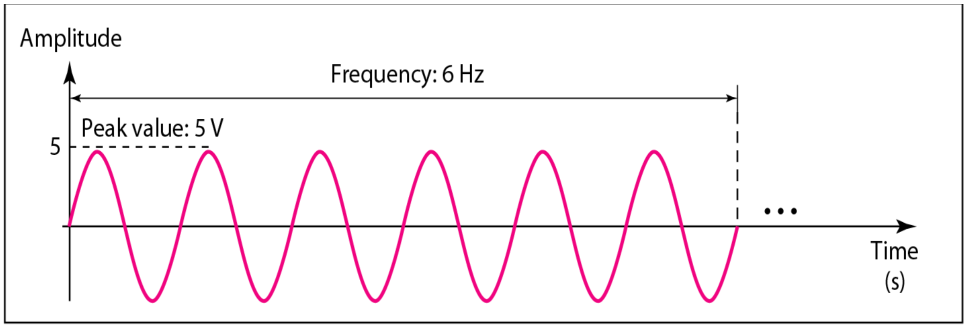
$$\lambda = \frac{c}{f}$$

Figure 3.6 Wave length and period

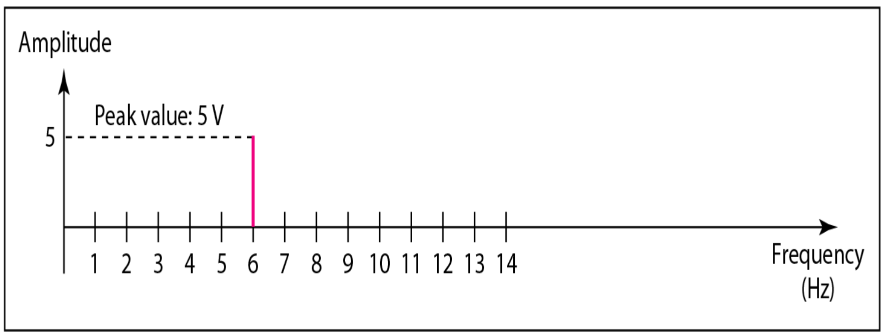
The propagation speed of electromagnetic signals depends on the medium and on the frequency of the signal. For example, in a vacuum, light is propagated with a speed of 3 x 10^8 m/s. That speed is lower in air and even lower in cable. The wavelength is normally measured in **micrometers (microns) instead of meters**.

TIME DOMAIN AND FREQUENCY DOMAIN

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



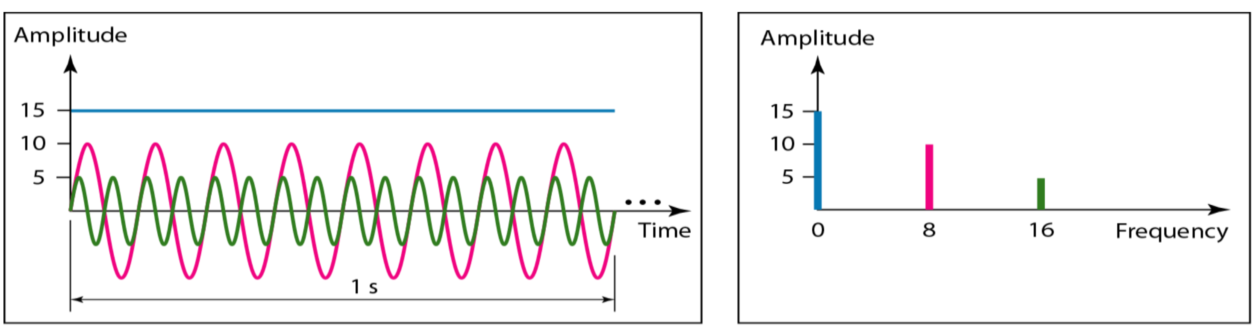
a. A sine wave in the time domain (peak value: 5 V, frequency: 6 Hz)



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

Figure 3.7 Time and frequency domain

For example, Below Figure 3.8 shows **thre**e sine waves, each with **different amplitude and frequency. All can be represented by three spikes in the frequency domain**.



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

b. Frequency-domain representation of the same three signals

Figure 3.8 Time and frequency domain for three sine waves

COMPOSITE SIGNAL: -

A signal composed of more than one sine wave. According to Fourier analysis, any **composite signal is a combination of simple sine waves with different frequencies, amplitudes, and phases**. If the composite signal is periodic, the decomposition gives a series of signals with **discrete frequencies**; if the composite signal is nonperiodic, the decomposition gives a combination of sine waves with **continuous frequencies** as shown below

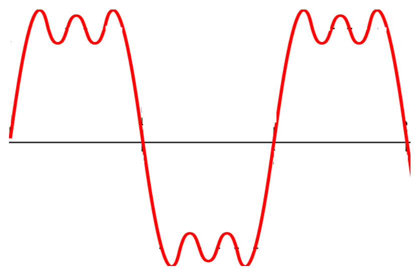
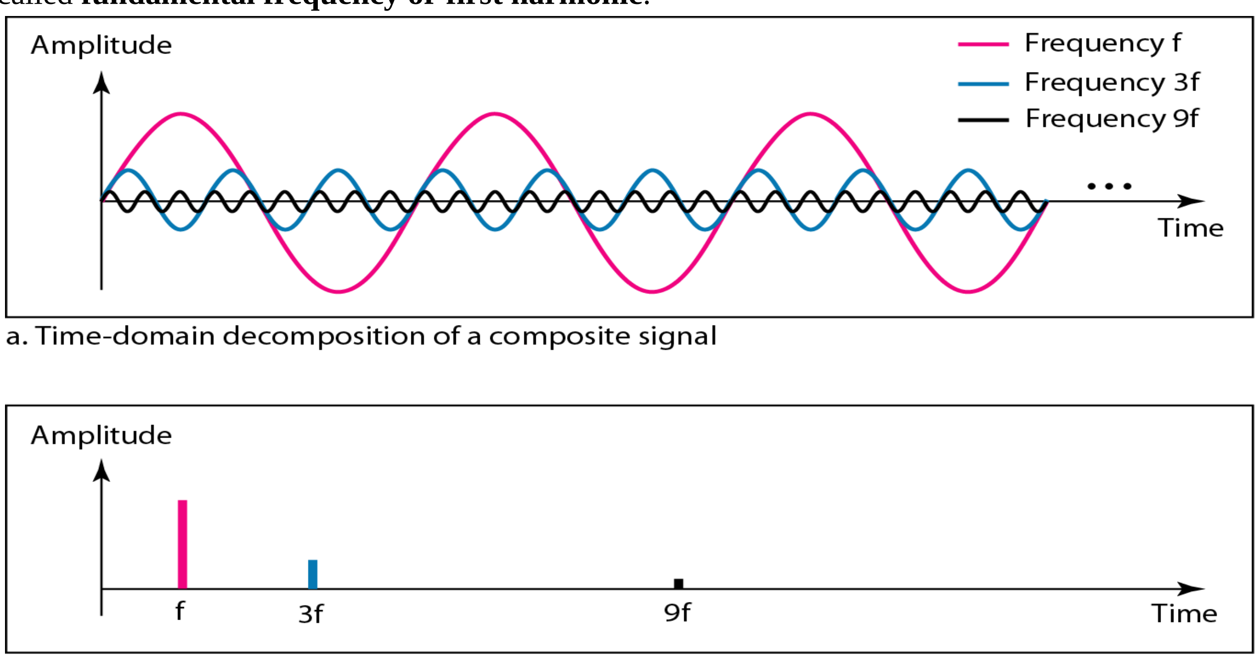


Figure 3.9: Composite periodc signal

The amplitude of the sine wave with frequency f is almost the same as the peak amplitude of the composite signal. As the frequency of the composite signal is same as the frequency of this signal so it is called **fundamental frequency or first harmonic**.



b. Frequency-domain decomposition of the composite signal

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

The amplitude of the sine wave with frequency 3f is one third of the first frequency. As the frequency is 3 times of the first frequency so it is 3rd harmonic. The amplitude of the sine wave with frequency 9f is one ninth of the first frequency. As the frequency is 9 times of the first frequency so it is 9th harmonic.

Example 3.9

Figure 3.11 shows a nonperiodic composite signal. It can be the signal created by a microphone or a telephone set when a word or two is pronounced.

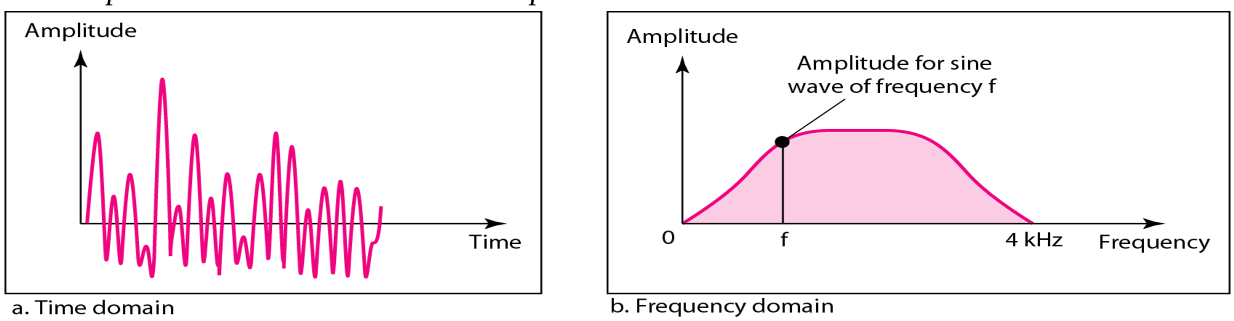
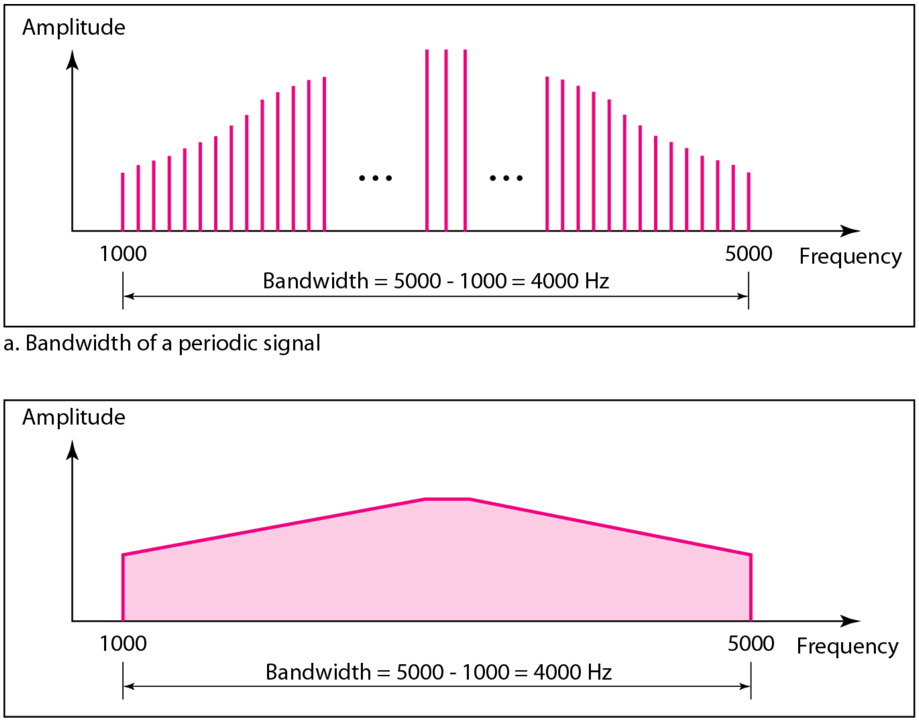


Figure 3.11 The *time and frequency domains of a nonperiodic signal*

BANDWIDTH: -

The range of frequencies contained in a composite signal. The bandwidth of a composite signal is the difference between the **highest and the lowest frequencies contained in that signal**. *Below figure shows bandwidth of periodic and non-periodic signals*



a. Bandwidth of a periodic signal

b. Bandwidth of a nonperiodic signal

figure 3.12: bandwidth of periodic and non periodic signals

Example 3.10

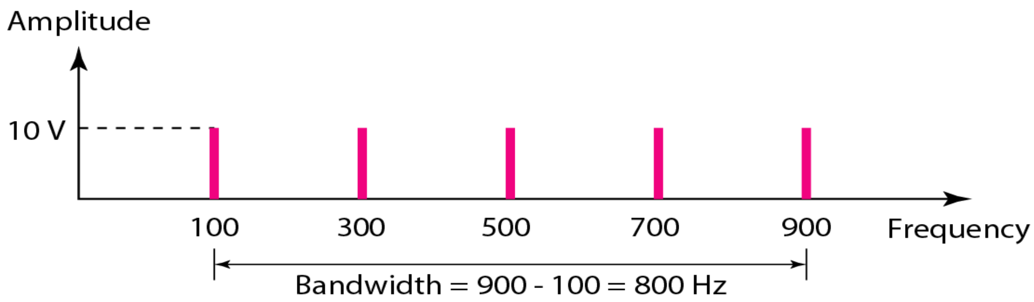
If a periodic signal is decomposed into five sine waves with frequencies of 100, 300, 500, 700, and 900 Hz, what is its bandwidth? Draw the spectrum, assuming all components have a maximum amplitude of 10 V.

Solution

Let f_h be the highest frequency, f_l the lowest frequency, and B the bandwidth. Then

$$\boxed{B = f_h - f_l = 900 - 100 = 800 \text{ Hz}}$$

The spectrum has only five spikes, at 100, 300, 500, 700, and 900 Hz .



Example 3.11

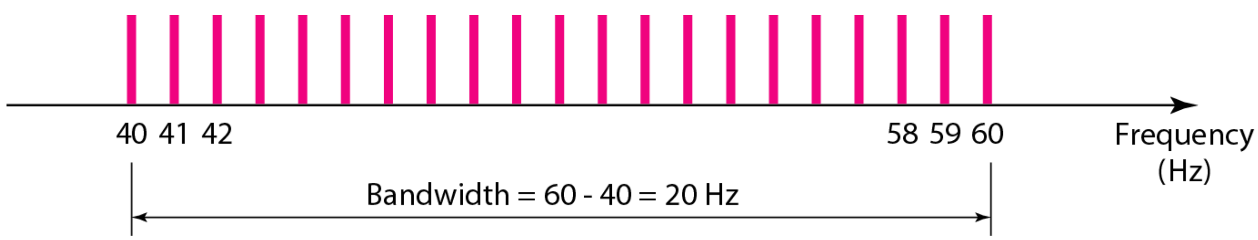
A periodic signal has a bandwidth of 20 Hz. The highest frequency is 60 Hz. What is the lowest frequency? Draw the spectrum if the signal contains all frequencies of the same amplitude.

Solution

Let f_h be the highest frequency, f_l the lowest frequency, and B the bandwidth. Then

$$\boxed{B = f_h - f_l \;\Rightarrow\; 20 = 60 - f_l \;\Rightarrow\; f_l = 60 - 20 = 40 \text{ Hz}}$$

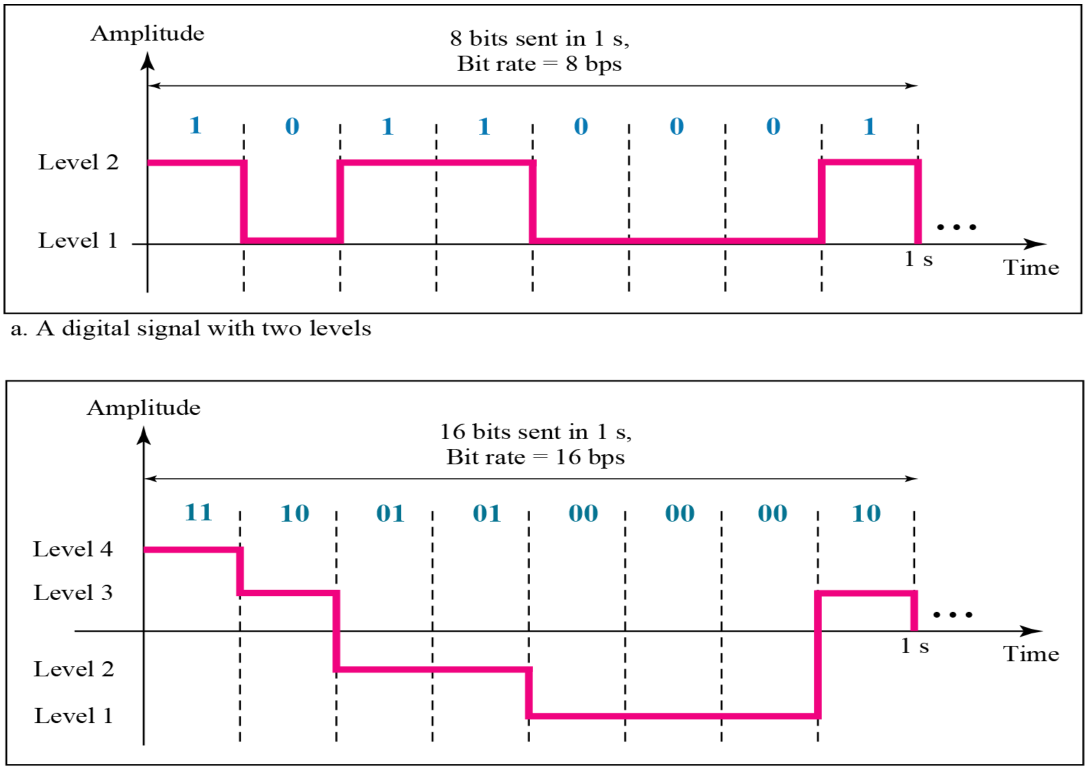
The spectrum contains all integer frequencies. We show this by a series of spikes.



DIGITAL SIGNALS

Information can also be represented by a digital signal. For example, a 1 can be encoded as a positive voltage and a 0 as zero voltage. A digital signal can have more than two levels. Below figure we are sending in two levels 1 bit and 2 bits in another.

In general, if signal has L levels, then each level needs: $\log_2 L$ *bits*



a. A digital signal with two levels

b. A digital signal with four levels

Figure 3.16 Two digital signals: one with two signal and the other with four signal levels

Example 3.16

A digital signal has eight levels. How many bits are needed per level? We calculate the number of bits from the formula:

$$\text{Number of bits per level} = \log_2 8 = 3$$

Each signal level is represented by 3 bits.

<u>The bit rate</u> is the number of bits sent in Is, expressed in bits per second (bps). Figure 3.16 shows the bit rate for two signals.

The bit length is the distance one bit occupies on the transmission medium.

Bit length = propagation speed * bit duration

TRANSMISSION OF DIGITAL SIGNALS

We can transmit a digital signal by using one of two different approaches:

1. ***Baseband transmission.***
2. ***Broadband transmission (using modulation).***

Baseband transmission means sending a digital signal over a channel without changing the digital signal to an analog signal.

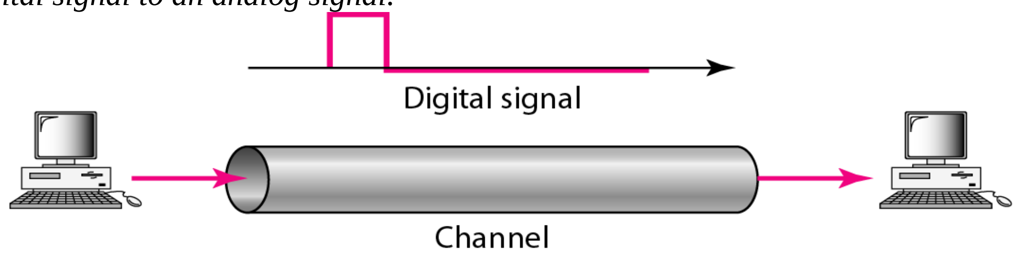


Figure 3.18 Baseband *transmission*

Baseband transmission requires that we have a low-pass channel, a channel with a bandwidth that **starts from zero**. This is the case if we have a dedicated medium with a bandwidth constituting only one channel. For example, the entire bandwidth of a cable **connecting two computers is one single channel**. As another example, we may connect several computers to a bus, but not allow more than two stations to communicate at a time. *a low-pass channel with infinite band-width is ideal, but we cannot have such a channel in real life.*

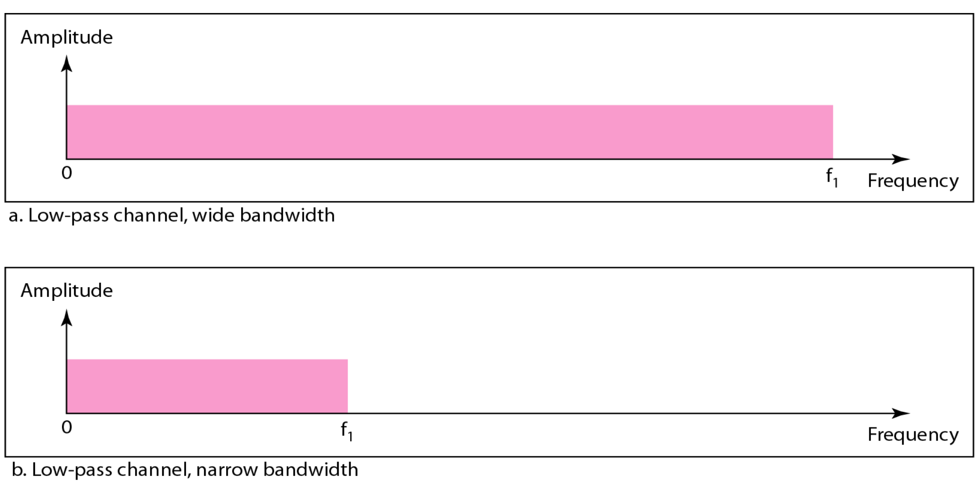


Figure 3.19 Bandwidth *of two low-pass channels*

Case 1: Low-Pass Channel with Wide Bandwidth

If we want to preserve the exact form of a non-periodic digital signal then we need to send the entire spectrum, the continuous range of frequencies between zero and infinity. This is possible if we have a dedicated medium with an infinite bandwidth between the sender and receiver that preserves the exact amplitude of each component of the composite signal. It is possible inside a computer but it is not possible between 2 devices.

This means that if we have a medium, such as a coaxial cable or fiber optic, with a very wide bandwidth, two stations can communicate by using digital signals with very good accuracy.

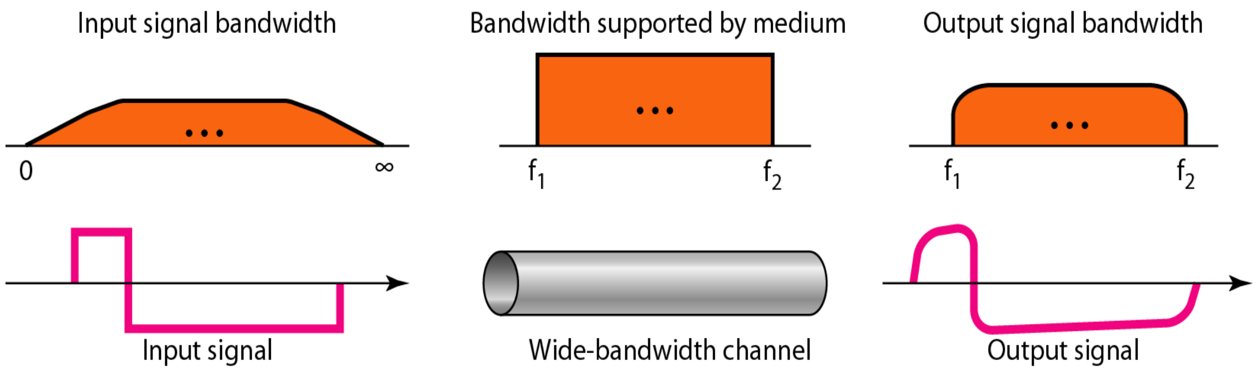


Figure 3.20: Baseband transmission using a dedicated medium

Example 3.22 What is the required bandwidth of a low-pass channel if we need to send 1 Mbps by using baseband transmission?

Solution: ***The answer depends on the accuracy desired.***

a. The minimum bandwidth, is B = bit rate /2, or 500 kHz.

b. A better solution is to use the first and the 3rd harmonics with B = 3 × 500 kHz = 1.5 MHz.

c. Still a better solution is to use the 1st, 3rd and 5th harmonics with B = 5 × 500kHz = 2.5 MHz.

Bit Rate	*Harmonic 1*	*Harmonics 1, 3*	*Harmonics 1, 3, 5*
$n = 1$ kbps	$B = 500$ Hz	$B = 1.5$ kHz	$B = 2.5$ kHz
$n = 10$ kbps	$B = 5$ kHz	$B = 15$ kHz	$B = 25$ kHz
$n = 100$ kbps	$B = 50$ kHz	$B = 150$ kHz	$B = 250$ kHz

Table 3.2: Bandwidth requirements

Example 3.22 We have a low-pass channel with bandwidth 100 kHz. What is the maximum bit rate of this channel?

Solution

The maximum bit rate can be achieved if we use the first harmonic. The bit rate is 2 times the available bandwidth, or 200 kbps.

Broadband transmission: -

If the available channel is a band pass channel, we cannot send the digital signal directly to the channel, **we need to convert the digital signal to an analog signal before transmission is as shown below**

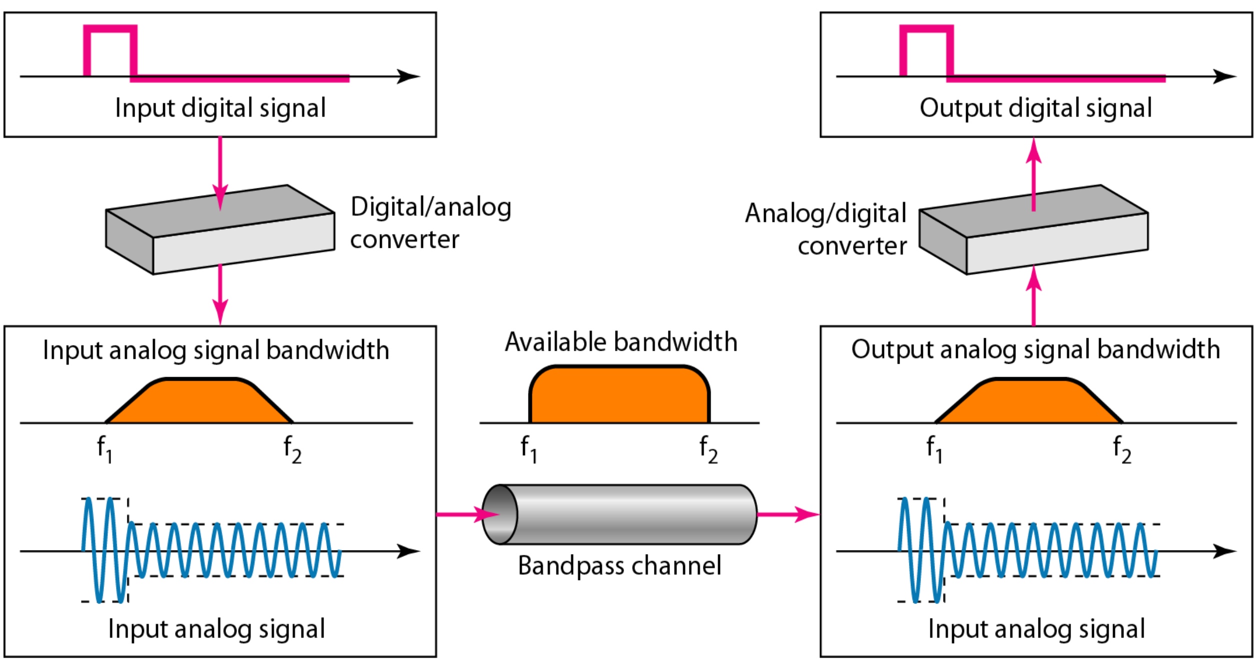


Figure 3.24 Modulation *of a digital signal for transmission on a band pass channel*

TRANSMISSION IMPAIRMENT

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

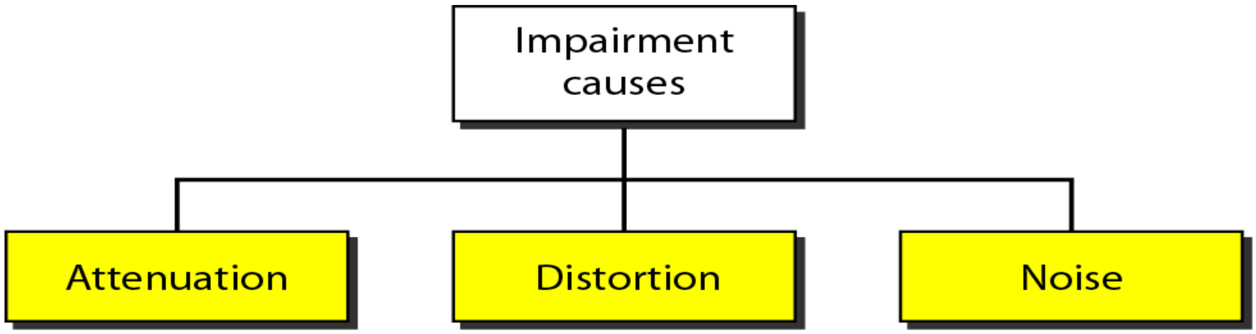


Figure 3.18 ***Causes of impairment***

ATTENUATION: -

Means a **loss of energy**. When a signal (simple or composite) travels through a medium, it loses some of its energy in overcoming the resistance of the medium. That is why a wire carrying electric **signals gets warm**. Some of the electrical energy in the signal is converted to heat. **To compensate for this loss, amplifiers are used to amplify the signal**. Figure 3.26 shows the effect of attenuation and amplification.

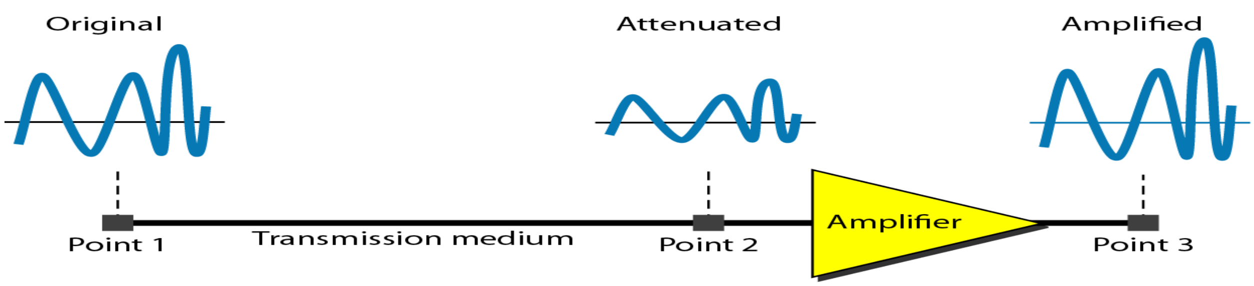


Figure 3.26 ***Attenuation***

Decibel: To measure signal gain or lost, uses decibels which measures the strength of signal at two different points. In general

$$dB = 10\log_{10}(p2/p1)$$

Where p1 and p2 are variables at point1 and point2.

Suppose a signal travels through a transmission medium and its power is reduced to one-half. This means that P_2 is (1/2) P1. In this case, the attenuation can be calculated as:

$$10 \log_{10} \frac{P_2}{P_1} = 10 \log_{10} \frac{0.5 P_1}{P_1} = 10 \log_{10} 0.5 = 10(-0.3) = -3 \text{ dB}$$

A loss of 3 dB (–3 dB) is equivalent to losing one-half the power.

Example 3.27

A signal travels through an amplifier, and its power is increased 10 times. This means that P_2 = $10P_1$. In this case, the amplification (gain of power) can be calculated as

$$10 \log_{10} \frac{P_2}{P_1} = 10 \log_{10} \frac{10 P_1}{P_1}$$

$$= 10 \log_{10} 10 = 10(1) = 10 \text{ dB}$$

Example 3.29

Sometimes the decibel is used to measure signal power in milliwatts. In this case, it is referred to as dB_m and is calculated as dB_m = 10 log10 P_m, where P_m is the power in milliwatts. Calculate the power of a signal with $dB_m = -30$.

Solution

We can calculate the power in the signal as

$$\text{dB}_\text{m} = 10 \log_{10} P_m = -30$$
$$\log_{10} P_m = -3 \qquad P_m = 10^{-3} \text{ mW}$$

DISTORTION: -

Means that the signal **changes its form or shape**. Distortion can occur in a composite signal made of **different frequencies**. Each signal component has its own propagation speed through a medium**. Differences in delay may create a difference in phase. The** signal components at the receiver have phases different from what they had at the sender. As shown below.

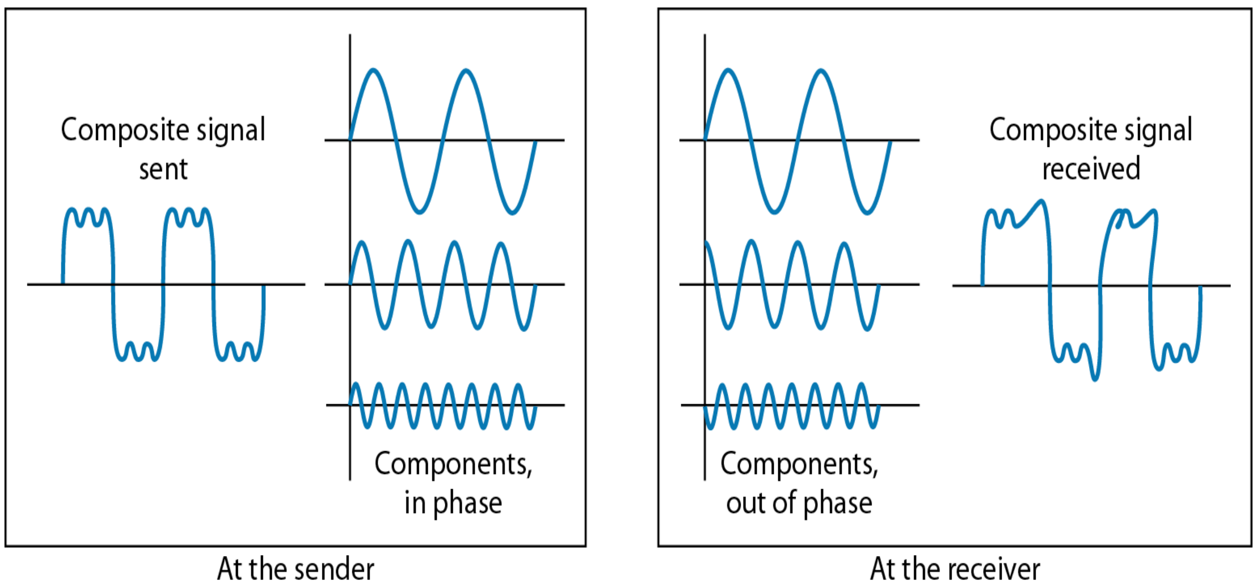


Figure 3.28 Distortion

NOISE;- ***I***s another cause of impairment. The following types of noise may corrupt the signal.

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

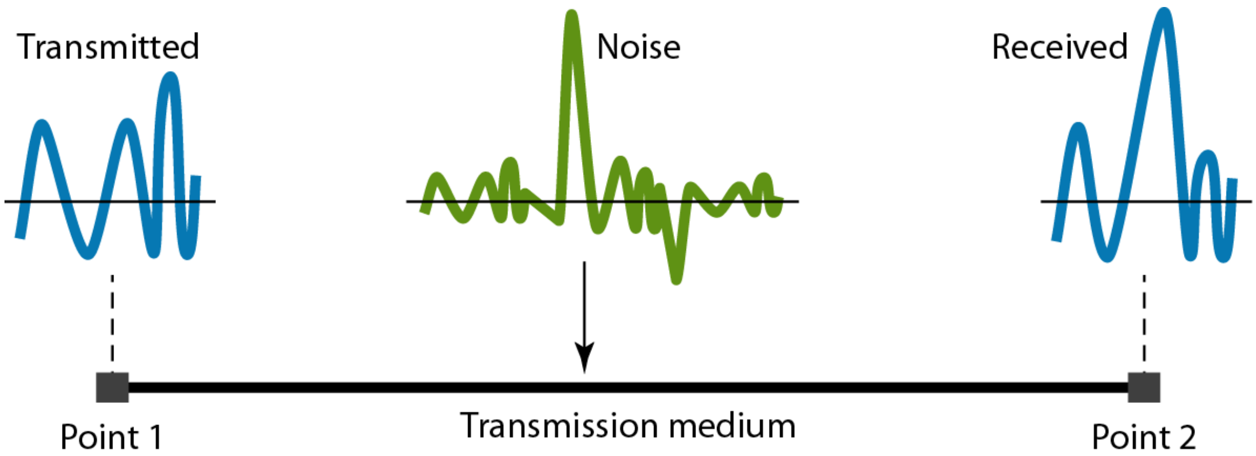


Figure 3.29 Noise

SIGNAL-TO-NOISE RATIO (SNR)

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

SNR = average signal power / average noise power

Consider the average signal power and the average noise power, these may change with time. **SNR is actually the ratio of what is wanted (signal) to what is not wanted (noise). A high SNR means the signal is less corrupted by noise; a low SNR means the signal is more corrupted by noise.** SNR is the ratio of two powers, it is often described in decibel units, SNRdB, defined as **SNR = 10 $\log_{10}$ SNR..** ***Below figure shows the idea of SNR.***

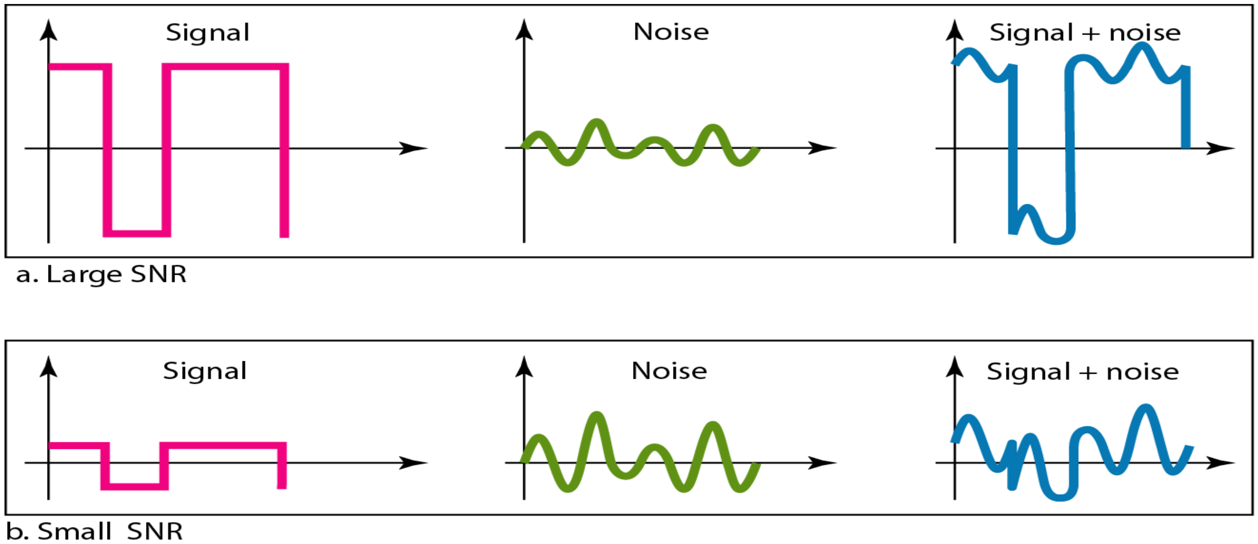


Figure 3.30 Two cases of SNR: a high SNR and a low SNR

Example: The power of a signal is 10 mW and the power of the noise is 1 μW; what are the values of SNR and SNR_{dB} ?

Solution

The values of SNR and SNR_{dB} can be calculated as follows:

$$\text{SNR} = \frac{10{,}000\ \mu\text{W}}{1\ \text{mW}} = 10{,}000$$

$$\text{SNR}_{\text{dB}} = 10 \log_{10} 10{,}000 = 10 \log_{10} 10^4 = 40$$

Example : The values of SNR and SNR_{dB} for a noiseless channel are

$$\text{SNR} = \frac{\text{signal power}}{0} = \infty$$

$$\text{SNR}_{\text{dB}} = 10 \log_{10} \infty = \infty$$

We can never achieve this ratio in real life; it is an ideal.

DATA RATE LIMITS

A very important consideration in data communications is how fast we can send data, in bits per second, over a channel. Data rate depends on three factors:

1. ***The bandwidth available***
2. ***The level of the signals we use***
3. ***The quality of the channel (the level of noise)***

Increasing the levels of a signal may reduce the reliability of the system. Two theoretical formulas were developed to calculate the data rate: one by **Nyquist for a noiseless channel,** another by **Shannon for a noisy channel**. **The Shannon capacity gives us the upper limit; the Nyquist formula tells us how many signal levels we need**.

Noiseless channel: **the Nyquist bit rate formula defines the theoretical maximum bit rate**

$$\textbf{Bit Rate} = 2 * \text{bandwidth} * \log_2 L$$

Whereas

Bit rate is bit rate in bits per second

Bandwidth is bandwidth of signal

L is number of signal levels

Noisy Channel: Shannon Capacity*:* In reality, we cannot have a noiseless channel; the channel is always noisy. In 1944, Claude Shannon introduced a formula, called the Shannon capacity, to determine the theoretical highest data rate for a noisy channel:

$$\textbf{Capacity} = \text{bandwidth} * \log_2 (1 + \text{SNR})$$

Whereas

Capacity is capacity of the channel in bits per second

Bandwidth is bandwidth of signal

SNR is Signal-to-Noise ratio.

Example: Consider a noiseless channel with a bandwidth of 3000 Hz transmitting a signal with two signal levels. The maximum bit rate can be calculated as

$$\boxed{\text{BitRate} = 2 \times 3000 \times \log_2 2 = 6000 \text{ bps}}$$

Example: *Consider the same noiseless channel transmitting a signal with four signal levels (for each level, we send 2 bits). The maximum bit rate can be calculated as*

$$\text{BitRate} = 2 \times 3000 \times \log_2 4 = 12{,}000 \text{ bps}$$

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

Solution

We can use the Nyquist formula as shown:

$$265{,}000 = 2 \times 20{,}000 \times \log_2 L$$
$$\log_2 L = 6.625 \qquad L = 2^{6.625} = 98.7 \text{ levels}$$

Example: *Consider an extremely noisy channel in which the value of the signal-to-noise ratio is almost zero. In other words, the noise is so strong that the signal is faint. For this channel the capacity C is calculated as*

$$C = B \log_2 (1 + \text{SNR}) = B \log_2 (1 + 0) = B \log_2 1 = B \times 0 = 0$$

This means that the capacity of this channel is zero regardless of the bandwidth. In other words, we cannot receive any data through this channel.

Example: *We can calculate the theoretical highest bit rate of a regular telephone line. A telephone line normally has a bandwidth of 3000. The signal-to-noise ratio is usually 3162. For this channel the capacity is calculated as*

$$C = B \log_2 (1 + \text{SNR}) = 3000 \log_2 (1 + 3162) = 3000 \log_2 3163$$
$$= 3000 \times 11.62 = 34{,}860 \text{ bps}$$

This means that the highest bit rate for a telephone line is 34.860 kbps. If we want to send data faster than this, we can either increase the bandwidth of the line or improve the signal-to-noise ratio.

Example: *The signal-to-noise ratio is often given in decibels. Assume that* SNR_{dB} *= 36 and the channel bandwidth is 2 MHz. The theoretical channel capacity can be calculated as*

$$SNR_{dB} = 10 \log_{10} SNR \rightarrow SNR = 10^{SNR_{dB}/10} \rightarrow SNR = 10^{3.6} = 3981$$
$$C = B \log_2 (1 + SNR) = 2 \times 10^6 \times \log_2 3982 = 24 \text{ Mbps}$$

Example: *For practical purposes, when the SNR is very high, we can assume that SNR + 1 is almost the same as SNR. In these cases, the theoretical channel capacity can be simplified to*

$$C = B \times \frac{SNR_{dB}}{3}$$

For example, we can calculate the theoretical capacity of the previous example as

$$C = 2 \text{ MHz} \times \frac{36}{3} = 24 \text{ Mbps}$$

Example: *We have a channel with a 1-MHz bandwidth. The SNR for this channel is 63. What are the appropriate bit rate and signal level?*

Solution

First, we use the Shannon formula to find the upper limit.

$$C = B \log_2 (1 + SNR) = 10^6 \log_2 (1 + 63) = 10^6 \log_2 64 = 6 \text{ Mbps}$$

The Shannon formula gives us 6 Mbps, the upper limit. For better performance we choose something lower, 4 Mbps, for example. Then we use the Nyquist formula to find the number of signal levels.

$$4 \text{ Mbps} = 2 \times 1 \text{ MHz} \times \log_2 L \rightarrow L = 4$$

PERFORMANCE: -

The performance is measured by **Bandwidth, Throughput, Latency (Delay) and Bandwidth-Delay Product**

Bandwidth: **In networking, we use the term bandwidth in two contexts.**

- **Bandwidth in hertz refers to the range of frequencies in a composite signal or the range of frequencies that a channel can pass.**
- **Bandwidth in bits per second refers to the speed of bit transmission in a channel or link.**

An increase in bandwidth in hertz means an increase in bandwidth in bits per second.

Example 3.44

A network with bandwidth of 10 Mbps can pass only an average of 12,000 frames per minute with each frame carrying an average of 10,000 bits. What is the throughput of this network?

Solution

We can calculate the throughput as

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

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

Throughput and latency:

The throughput is a measure of how fast we can actually send data through a network. **Bandwidth and throughput is not same**. For example a link with bandwidth 1 Mbps but can handle only 200 kbps so throughput will be 200 kbps.

The latency or delay defines how long it takes for an entire message to completely arrive at the destination from the first bit is sent out from the source.

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

Propagation time measures the time required for a bit to travel from the source to the destination.

Transmission time is the time between the first bit leaving the sender and the last bit arriving at the receiver.

Queuing time is the required time for each intermediate or end device to hold the message before it can be processed.

Bandwidth-delay product defines the number of bits that can fill the link.

Bandwidth-delay product = bandwidth x delay

Example 3.45

What is the propagation time if the distance between the two points is 12,000 km? Assume the propagation speed to be 2.4 × 108 m/s in cable.

Solution

We can calculate the propagation time as

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

The example shows that a bit can go over the Atlantic Ocean in only 50 ms if there is a direct cable between the source and the destination.

Example 3.46

What are the propagation time and the transmission time for a 2.5-kbyte message (an e-mail) if the bandwidth of the network is 1 Gbps? Assume that the distance between the sender and the receiver is 12,000 km and that light travels at 2.4 × 108 m/s.

Solution

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

$$\text{Transmission time} = \frac{2500 \times 8}{10^{9}} = 0.020 \text{ ms}$$

Note that in this case, because the message is short and the bandwidth is high, the dominant factor is the propagation time, not the transmission time. The transmission time can be ignored

Example 3.47

What are the propagation time and the transmission time for a 5-Mbyte message (an image) if the bandwidth of the network is 1 Mbps? Assume that the distance between the sender and the receiver is 12,000 km and that light travels at 2.4×10^8 m/s.

Solution

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

$$\text{Transmission time} = \frac{5{,}000{,}000 \times 8}{10^6} = 40 \text{ s}$$

Note that in this case, because the message is very long and the bandwidth is not very high, the dominant factor is the transmission time, not the propagation time. The propagation time can be ignored

Transmission Time:

In data communications we don't send just 1 bit, we send a message. The first bit may take a time equal to the propagation time to reach its destination; the last bit also may take the same amount of time. However, there is a time between the first bit leaving the sender and the last bit arriving at the receiver. The first bit leaves earlier and arrives earlier; the last bit leaves later and arrives later. **The time required for transmission of a message depends on the size of the message and the bandwidth of the channel t**hen we hvae

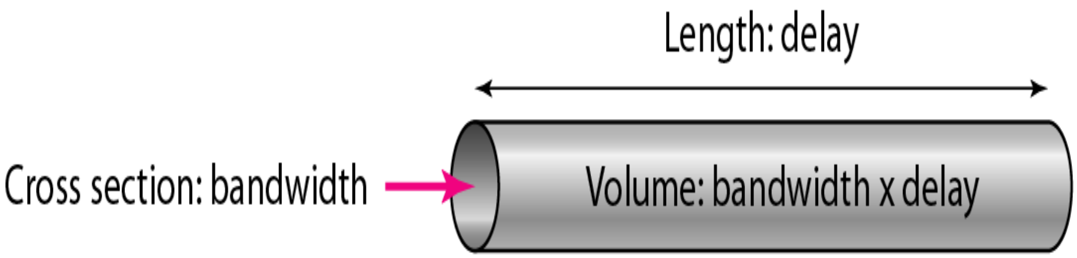
$$\textbf{Transmission time} = \frac{\textbf{Message size}}{\textbf{Bandwidth}}$$

Queuing Time:

Is **the time needed for each intermediate or end device to hold the message before it can be processed.** The queuing time is not a fixed factor; **it changes with the load imposed on the network**. When there is heavy traffic on the network, the queuing time increases.

Bandwidth-Delay Product:

Bandwidth and delay are two performance metrics of a link. Very important in data communications is the product of the two, the **bandwidth-delay product**.



Case 1. Let us assume that we have a link with a bandwidth of 1 bps Also assume that the delay of the link is 5 s we want to see what the bandwidth-delay product means in this case. From the figure, it can be said that this product 1 x 5 is the maximum number of bits that can fill the link. There can be no more than 5 bits at any time on the link as shown below

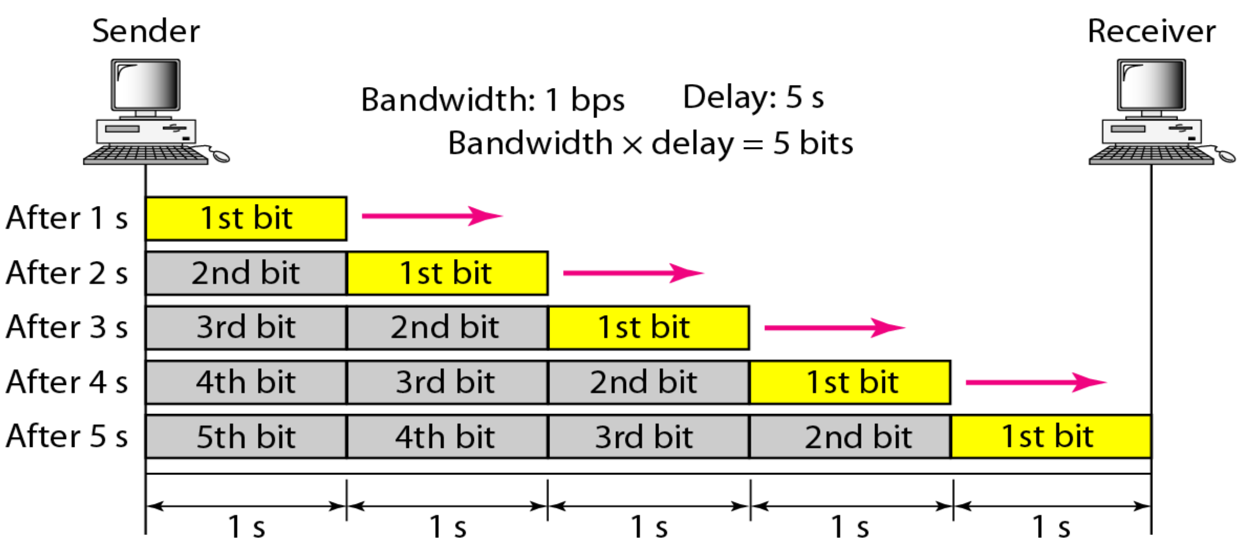


Figure 3.31: Filling link with bits(case 1)

Case 2. Now assume we have a bandwidth of 4 bps. The figure shows that there can be maximum 4 x 5 = 20 bits on the line. The reason is that, at each second, there are 4 bits on the line; the duration of each bit is 0.25s.as shown below

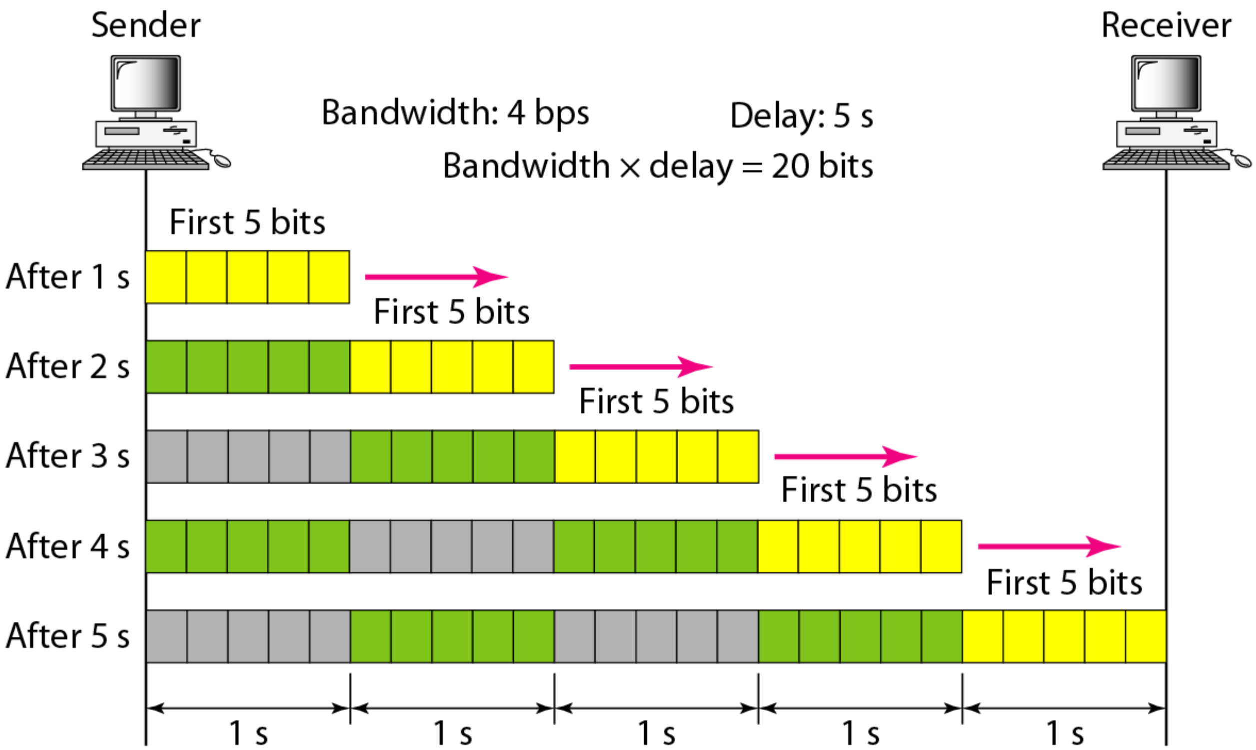


Figure 3.32: Filling link with bits(case 2)

The above two cases show that the product of bandwidth and delay is the number of bits that can fill the link. This measurement is important if we need to send data in bursts and wait for the acknowledgment of each burst before sending the next one.

To use the maximum capability of the link, we need to make the size of our burst 2 times the product of bandwidth and delay; we need to fill up the full-duplex channel (two directions). The sender should send a burst of data of (2 x bandwidth x delay) bits.

The sender then waits for receiver acknowledgment for part of the burst before sending another burst. The amount 2 x bandwidth x delay is the number of bits that can be in transition at any time.

Jitter

Time difference in packet inter-arrival time to their destination can be called **jitter**. Jitter is specific issue that normally exists in packet networks.