Introduction to Data Communications:

In Data Communications, *data* generally are defined as information that is stored in digital form. *Data communications* is the process of transferring digital information between two or more points. *Information* is defined as the knowledge or intelligence. Data communications can be summarized as the transmission, reception, and processing of digital information. For data communications to occur, the communicating devices must be part of a communication system made up of a combination of hardware (physical equipment) and software (programs). The effectiveness of a data communications system depends on four fundamental characteristics: delivery, accuracy, timeliness, and jitter.

A data communications system has five components:

1. **Message:** The message is the information (data) to be communicated. Popular forms of information include text, numbers, pictures, audio, and video.
2. **Sender:** The sender is the device that sends the data message. It can be a computer, workstation, telephone handset, video camera, and so on.
3. **Receiver:** The receiver is the device that receives the message. It can be a computer, workstation, telephone handset, television, and so on.
4. **Transmission medium:** The transmission medium is the physical path by which a message travels from sender to receiver. Some examples of transmission media include twisted-pair wire, coaxial cable, fiber-optic cable, and radio waves.
5. **Protocol**: A protocol is a set of rules that govern data communications. It represents an agreement between the communicating devices.

Standards Organizations for Data Communications

An association of organizations, governments, manufacturers and users form the standards organizations and are responsible for developing, coordinating and maintaining the standards. The intent is that all data communications equipment manufacturers and users comply with these standards. The primary standards organizations for data communication are:

1. **International Standard Organization (ISO)**

   ISO is the international organization for standardization on a wide range of subjects. It is comprised mainly of members from the standards committee of various governments throughout the world. It is even responsible for developing models which provides high level of system compatibility, quality enhancement, improved productivity and reduced costs. The ISO is also responsible for endorsing and coordinating the work of the other standards organizations.

2. **International Telecommunications Union-Telecommunication Sector (ITU-T)**

   ITU-T is one of the four permanent parts of the International Telecommunications Union based in Geneva, Switzerland. It has developed three sets of specifications: the *V series* for modem interfacing and data transmission over telephone lines, the *X series* for data transmission over public digital networks, email and directory services; the *I and Q series* for Integrated Services Digital Network (ISDN) and its extension Broadband ISDN. ITU-T membership consists of government authorities and representatives from many countries and it is the present standards organization for the United Nations.

3. Institute of Electrical and Electronics Engineers (IEEE)

IEEE is an international professional organization founded in United States and is compromised of electronics, computer and communications engineers. It is currently the world's largest professional society with over 200,000 members. It develops communication and information processing standards with the underlying goal of advancing theory, creativity, and product quality in any field related to electrical engineering.

4. American National Standards Institute (ANSI)

ANSI is the official standards agency for the United States and is the U.S voting representative for the ISO. ANSI is a completely private, non-profit organization comprised of equipment manufacturers and users of data processing equipment and services. ANSI membership is comprised of people form professional societies, industry associations, governmental and regulatory bodies, and consumer goods.

5. Electronics Industry Association (EIA)

EIA is a non-profit U.S. trade association that establishes and recommends industrial standards. EIA activities include standards development, increasing public awareness, and lobbying and it is responsible for developing the RS (recommended standard) series of standards for data and communications.

6. Telecommunications Industry Association (TIA)

TIA is the leading trade association in the communications and information technology industry. It facilitates business development opportunities through market development, trade promotion, trade shows, and standards development. It represents manufacturers of communications and information technology products and also facilitates the convergence of new communications networks.

7. Internet Architecture Board (IAB)

IAB earlier known as Internet Activities Board is a committee created by ARPA (Advanced Research Projects Agency) so as to analyze the activities of ARPANET whose purpose is to accelerate the advancement of technologies useful for U.S military. IAB is a technical advisory group of the Internet Society and its responsibilities are:

I. Oversees the architecture protocols and procedures used by the Internet.

II. Manages the processes used to create Internet Standards and also serves as an appeal board for complaints regarding improper execution of standardization process.

III. Responsible for administration of the various Internet assigned numbers

IV. Acts as a representative for Internet Society interest in liaison relationships with other organizations.

V. Acts as a source of advice and guidance to the board of trustees and officers of Internet Society concerning various aspects of internet and its technologies.

8. **Internet Engineering Task Force (IETF)**

   The IETF is a large international community of network designers, operators, vendors and researchers concerned with the evolution of the Internet architecture and smooth operation of the Internet.

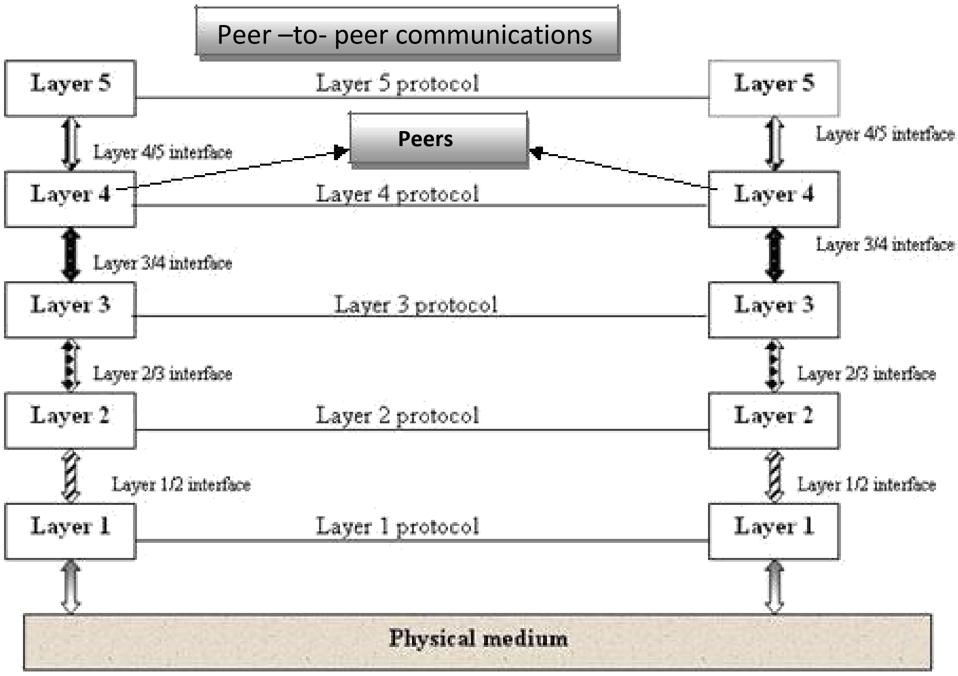
9. **Internet Research Task Force (IRTF)**

   The IRTF promotes research of importance to the evolution of the future Internet by creating focused, long-term and small research groups working on topics related to Internet protocols, applications, architecture and technology.

Layered Network Architecture

To reduce the design complexity, most of the networks are organized as a series of **layers** or **levels**, each one build upon one below it. The basic idea of a layered architecture is *to divide the design into small pieces*. Each layer adds to the services provided by the lower layers in such a manner that the highest layer is provided a full set of services to manage communications and run the applications. The benefits of the layered models are modularity and clear interfaces, i.e. open architecture and comparability between the different providers' components. A basic principle is to ensure independence of layers by defining services provided by each layer to the next higher layer without defining how the services are to be performed. This permits changes in a layer without affecting other layers. The basic elements of a layered model are services, protocols and interfaces. A ***service*** is a set of actions that a layer offers to another (higher) layer. ***Protocol*** is a set of rules that a layer uses to exchange information with a peer entity. These rules concern both the contents and the order of the messages used. Between the layers service interfaces are defined. The messages from one layer to another are sent through those interfaces.

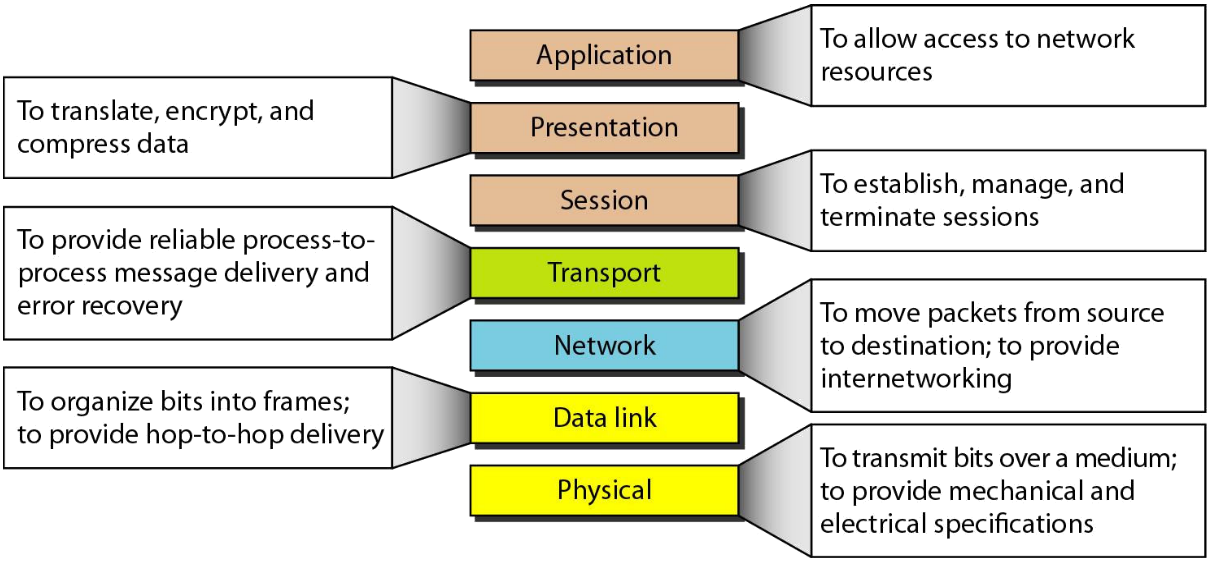
In a *n-layer* architecture, layer n on one machine carries on conversation with the layer n on other machine. The rules and conventions used in this conversation are collectively known as the *layer-n protocol*. Basically, a protocol is an agreement between the communicating parties on how communication is to proceed. Five-layer architecture is shown below**;** the entities comprising the corresponding layers on different machines are called ***peers***. In other words, it is the peers that communicate using protocols. In reality, no data is transferred from layer n on one machine to layer n of another machine. Instead, each layer passes data and control information to the layer immediately below it, until the lowest layer is reached. Below layer-1 is the physical layer through which actual communication occurs.



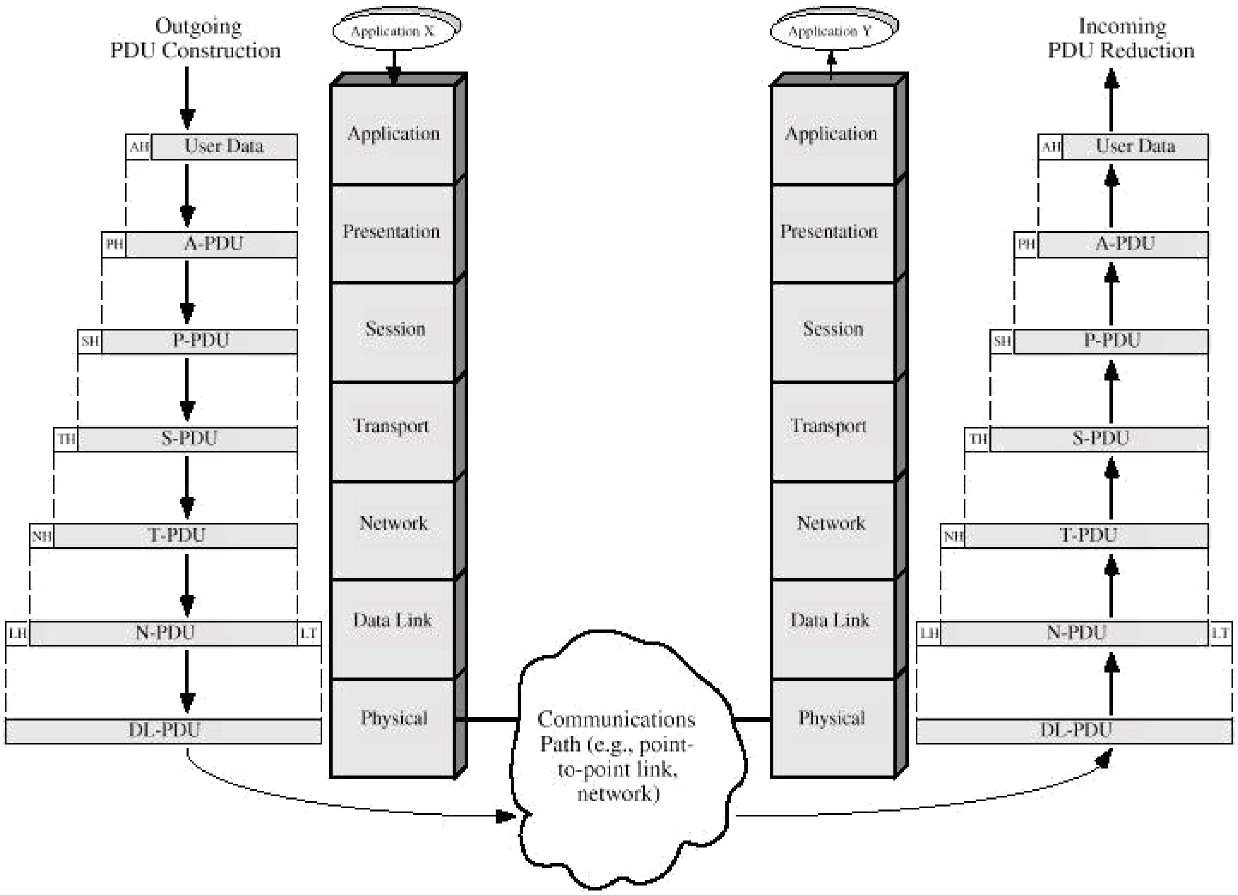
With layered architectures, communications between two corresponding layers requires a unit of data called a ***protocol data unit (PDU)***. A PDU can be a header added at the beginning of a message or a trailer appended to the end of a message. Data flows downward through the layers in the source system and upwards at the destination address. As data passes from one layer into another, headers and trailers are added and removed from the PDU. This process of adding or removing PDU information is called ***encapsulation/decapsulation.*** Between each pair of adjacent layers there is an ***interface***. The *interface* defines which primitives operations and services the lower layer offers to the upper layer adjacent to it. A set of layers and protocols is known as **network architecture**. A list of protocols used by a certain system, one protocol per layer, is called **protocol stack**.

Open Systems Interconnection (OSI)

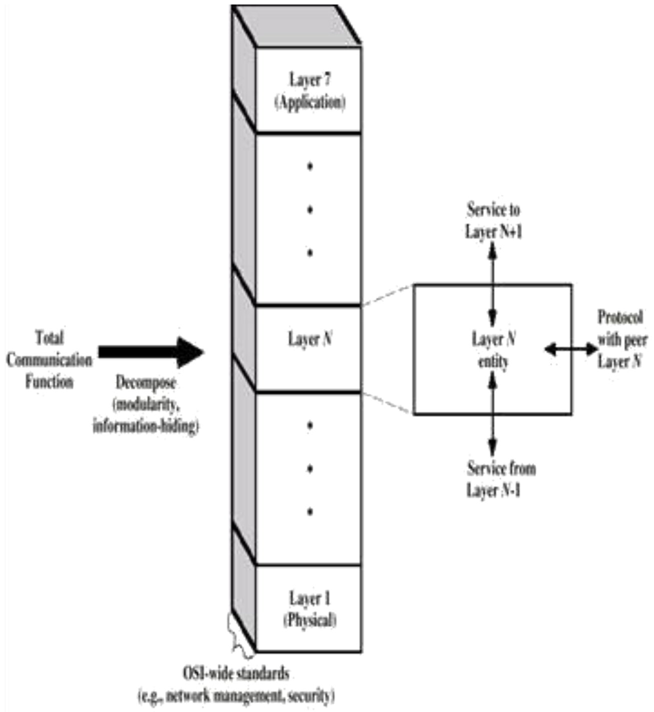
International standard organization (ISO) established a committee in 1977 to develop architecture for computer communication and the OSI model is the result of this effort. In 1984, the Open Systems Interconnection (OSI) reference model was approved as an international standard for communications architecture. The term "*open*" denotes the ability to connect any two systems which conform to the reference model and associated standards. The OSI model describes how information or data makes its way from application programmes (such as spreadsheets) through a network medium (such as wire) to another application programme located on another network. The OSI reference model divides the problem of moving information between computers over a network medium into **SEVEN** smaller and more manageable problems. The seven layers are:



The lower 4 layers (transport, network, data link and physical —Layers 4, 3, 2, and 1) are concerned with the flow of data from end to end through the network. The upper four layers of the OSI model (application, presentation and session—Layers 7, 6 and 5) are orientated more toward services to the applications. Data is Encapsulated with the necessary protocol information as it moves down the layers before network transit.



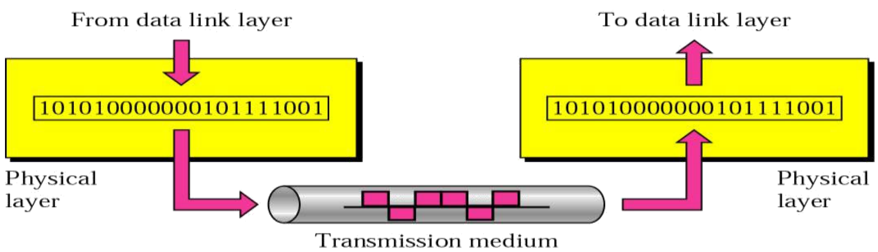
As with any layered architecture, overhead information is added to a PDU in the form of headers and trailers. Each layer provides a service to the layer above it in the protocol specification. Each layer communicates with the same layer's software or hardware on other computers.



Layer	Example
7.) Application	HTTP, FTP, SMTP
6.) Presentation	ASCII, JPEG, PGP
5.) Session	BOOTP, NetBIOS, DHCP, DNS
4.) Transport	TCP, UDP, SPX
3.) Network	IP, IPX, ICMP
2.) Data Link	Ethernet, Token Ring, Frame Relay
1.) Physical	Bits, Interfaces, Hubs

Physical Layer {***the physical layer is responsible for transmitting individual bits from one node to the next***}

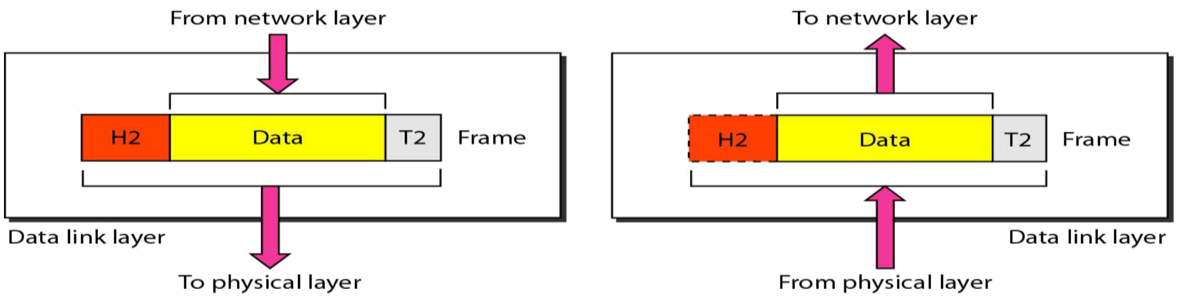
The physical layer is the lowest layer of the OSI hierarchy and coordinates the functions required to transmit a bit stream over a physical medium. It also defines the procedures and functions that physical devices and interfaces have to perform for transmission occur. The physical layer specifies the type of transmission medium and the transmission mode (simplex, half duplex or full duplex) and the physical, electrical, functional and procedural standards for accessing data communication networks.



Transmission media defined by the physical layer include metallic cable, optical fiber cable or wireless radio-wave propagation. The physical layer also includes the *carrier system* used to propagate the data signals between points in the network. The carrier systems are simply communication systems that carry data through a system using either metallic or optical fiber cables or wireless arrangements such as microwave, satellites and cellular radio systems.

Data-link Layer {***the data link layer is responsible for transmitting frames from one node to the next***}

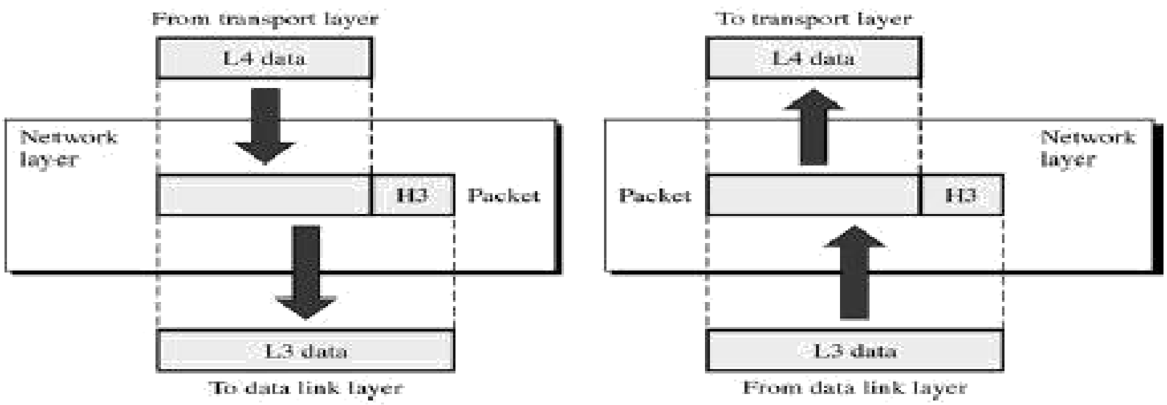
The data link layer transforms the physical layer, a raw transmission facility, to a reliable link and is responsible for node-to-node delivery. It makes the physical layer appear error free to the upper layer (network layer).



The data link layer packages data from the physical layer into groups called blocks, frames or packets. 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 physical address of the sender (source address) and/or receiver (destination address) of the frame. The data-link layer provides flow-control, access-control, and error-control.

Network Layer **{is responsible for the delivery of individual packets from the source host to the destination host}**

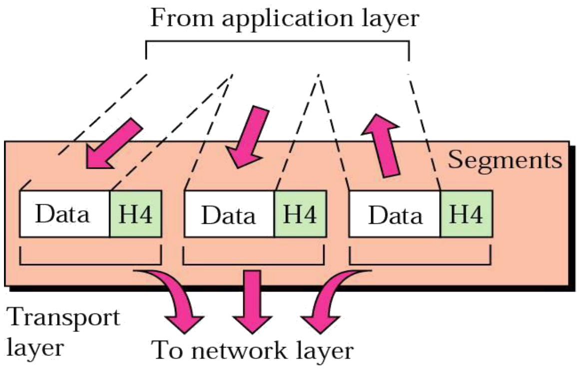
The network layer provides details that enable data to be routed between devices in an environment using multiple networks, subnetworks or both. This is responsible for addressing messages and data so they are sent to the correct destination, and for translating logical addresses and names (like a machine name FLAME) into physical addresses. This layer is also responsible for finding a path through the network to the destination computer.



The network layer provides the upper layers of the hierarchy with independence from the data transmission and switching technologies used to interconnect systems. Networking components that operate at the network layer include routers and their software.

Transport Layer *{is responsible for delivery of a message from one process to another}*

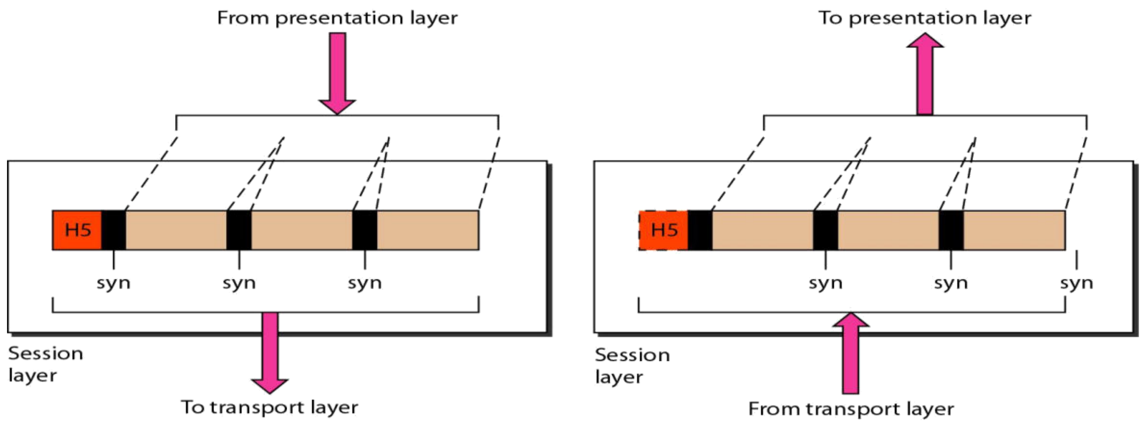
The transport layer controls and ensures the end-to-end integrity of the data message propagated through the network between two devices, providing the reliable, transparent transfer of data between two endpoints.



Transport layer responsibilites includes message routing, segmenting, error recovery and two types of basic services to an upper-layer protocol: connection oriented and connectionless. The transport layer is the highest layer in the OSI hierarchy in terms of communicatons and may provide data tracking, connection flow control, sequencing of data, error checking, and application addressing and identification.

Session Layer {**responsible for dialog control and synchronization**}

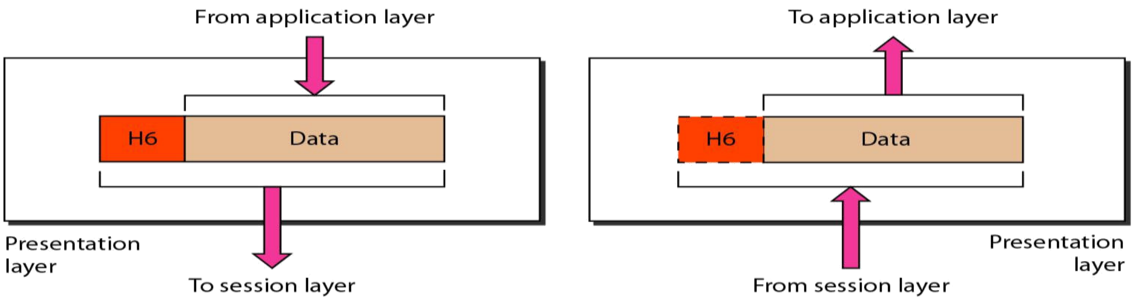
Session layer, some times called the dialog controller provides mechanism for controlling the dialogue between the two end systems. It defines how to start, control and end conversations (called sessions) between applications.



Session layer protocols provide the logical connection entities at the application layer. These applications include file transfer protocols and sending email. Session responsibilities include network log-on and log-off procedures and user authentication. Session layer characteristics include virtual connections between applications, entities, synchronization of data flow for recovery purposes, creation of dialogue units and activity units, connection parameter negotiation, and partitioning services into functional groups.

Presentation Layer {**responsible for translation, compression, and encryption**}

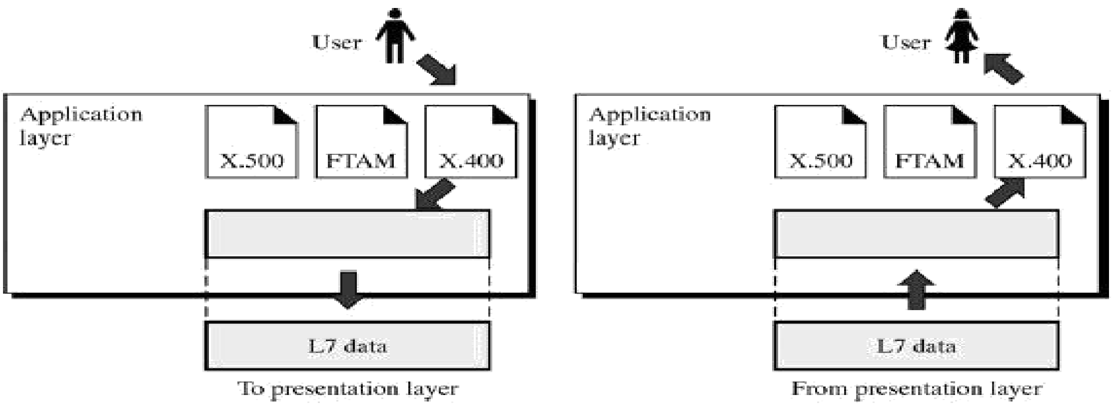
The presentation layer provides independence to the application processes by addressing any code or syntax conversion necessary to present the data to the network in a common communications format. It specifies how end-user applications should format the data.



The presentation layer translated between different data formats and protocols. Presentation functions include data file formatting, encoding, encryption and decryption of data messages, dialogue procedures, data compression algorithms, synchronization, interruption, and termination.

Application Layer {**responsible for providing services to the user**}

The application layer is the highest layer in the hierarchy and is analogous to the general manager of the network by providing access to the OSI environment. The applications layer provides distributed information services and controls the sequence of activities within and application and also the sequence of events between the computer application and the user of another application.



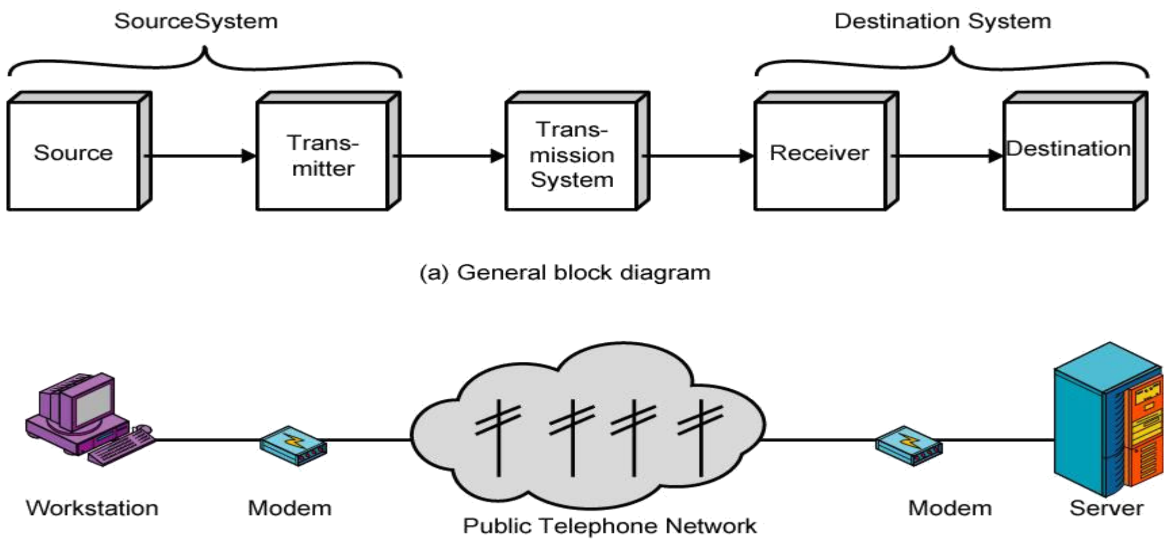
The application layer communicates directly with the user's application program. User application processes require application layer service elements to access the networking environment. The service elements are of two types: CASEs *(common application service elements)* satisfying particular needs of application processes like association control, concurrence and recovery. The second type is SASE (*specific application service elements)* which include TCP/IP stack, FTP, SNMP, Telnet and SMTP.

Data Communication Circuits

The underlying purpose of a digital communications circuit is to provide a transmission path between locations and to transfer digital information from one station (node, where computers or other digital equipment are located) to another using electronic circuits. Data communications circuits utilize electronic communications equipment and facilities to interconnect digital computer equipment. Communication facilities are physical means of interconnecting stations and are provided to data communications users through public telephone networks (PTN), public data networks (PDN), and a multitude of private data communications systems.

The following figure shows a simple two-station data communications circuit. The main components are:

Source: - This device generates the data to be transmitted; examples are mainframe computer, personal computer, workstation etc. The source equipment provides a means for humans to enter data into system.



(b) Example

Transmitter: - A transmitter transforms and encodes the information in such a way as to produce electromagnetic signals that can be transmitted across some sort of transmission system. For example, a modem takes a digital bit stream from an attached device such as a personal computer and transforms that bit stream into an analog signal that can be handled by the telephone network.

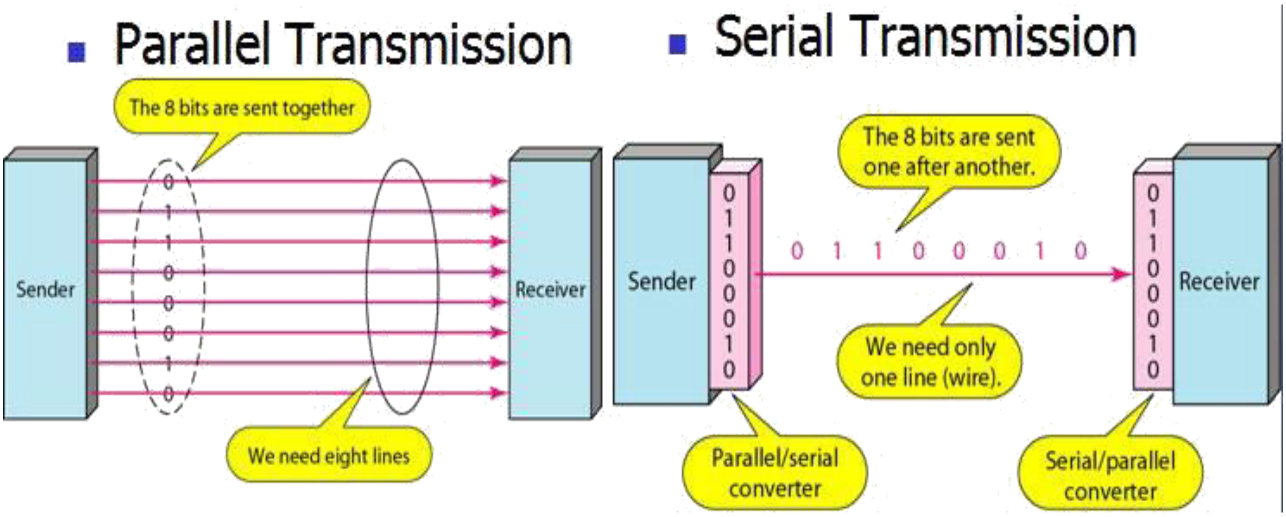
Transmission medium: - The transmission medium carries the encoded signals from the transmitter to the receiver. Different types of transmission media include free-space radio transmission (i.e. all forms of wireless transmission) and physical facilities such as metallic and optical fiber cables.

Receiver: - The receiver accepts the signal from the transmission medium and converts it into a form that can be handled by the destination device. For example, a modem will accept an analog signal coming from a network or transmission line and convert it into a digital bit stream.

Destination: - Takes the incoming data from the receiver and can be any kind of digital equipment like the source.

Serial and Parallel Data Transmission

There are two methods of transmitting digital data namely ***parallel and serial*** transmissions. In parallel data transmission, all bits of the binary data are transmitted simultaneously. For example, to transmit an 8-bit binary number in parallel from one unit to another, eight transmission lines are required. Each bit requires its own separate data path. All bits of a word are transmitted at the same time. This method of transmission can move a significant amount of data in a given period of time. Its disadvantage is the large number of interconnecting cables between the two units. For large binary words, cabling becomes complex and expensive. This is particularly true if the distance between the two units is great. Long multiwire cables are not only expensive, but also require special interfacing to minimize noise and distortion problems. Serial data transmission is the process of transmitting binary words a bit at a time. Since the bits time-share the transmission medium, only one interconnecting lead is required.



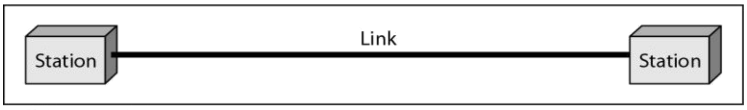
While serial data transmission is much simpler and less expensive because of the use of a single interconnecting line, it is a very slow method of data transmission. Serial data transmission is useful in systems where high speed is not a requirement. Parallel communication is used for short-distance data communications and within a computer, and serial transmission is used for long-distance data communications.

Data Communication Circuit Arrangements

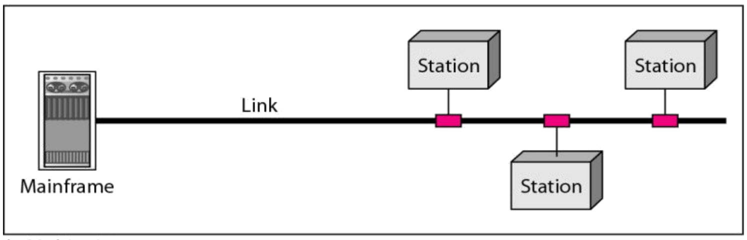
A data communications circuit can be described in terms of circuit configuration and transmission mode.

Circuit Configurations

Data communications networks can be generally categorized as either two point or multipoint. A *two-point* configuration involves only two locations or stations, whereas a *multipoint* configuration involves three or more stations.



a. Point-to-point

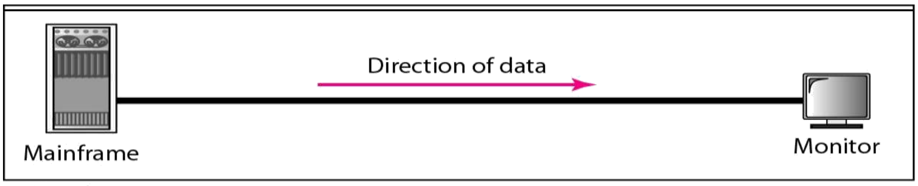


b. Multipoint

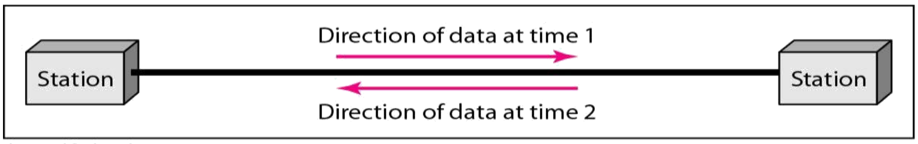
A two-point circuit involves the transfer of digital information between a mainframe computer and a personal computer, two mainframe computers or two data communications networks. A multi-point network is generally used to interconnect a single mainframe computer (host) to many personal computers or to interconnect many personal computers and capacity of the channel is either *Spatially shared*: Devices can use the link simultaneously or *Timeshare*: Users take turns

Transmission Modes

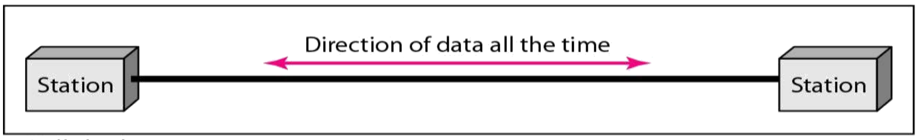
There are four modes of transmission for data communications circuits:



a. Simplex



b. Half-duplex



c. Full-duplex

In ***simplex mode(SX)***, the communication is unidirectional, as on a one-way street. Only one of the two devices on a link can transmit; the other can only receive. Commercial radio broadcasting is an example. Simplex lines are also called receive-only, transmit-only or one-way-only lines.

In ***half-duplex(HDX)*** mode, each station can both transmit and receive, but not at the same time. When one device is sending, the other can only receive, and vice versa. The half-duplex mode is used in cases where there is no need for communication in both directions at the same time; the entire capacity of the channel can be utilized for each direction. Citizens band (CB) radio is an example where push to talk (PTT) is to be pressed or depressed while sending and transmitting.

In ***full-duplex mode(FDX)*** (called duplex), both stations can transmit and receive simultaneously. One common example of full-duplex communication is the telephone network. The full-duplex mode is used when communication in both directions is required all the time. The capacity of the channel must be divided between the two directions.

In ***full/full duplex (F/FDX)*** mode, transmission is possible in both directions at the same time but not between the same two stations (i.e. station 1 transmitting to station 2, while receiving from station 3). F/FDX is possible only on multipoint circuits. Postal system can be given as a person can be sending a letter to one address and receive a letter from another address at the same time.

Data Communications Networks

Any group of computers connected together can be called a *data communications network,* and the process of sharing resources between computers over a data communications network is called *networking*. The most important considerations of a data communications network are *performance, transmission rate, reliability and security.*

Network Components, Functions, and Features

The major components of a network are end stations, applications and a network that will support traffic between the end stations. Computer networks all share common devices, functions, and features, including servers, clients, transmission media, shared data, shared printers and other peripherals, hardware and software resources, network interface card (NIC), local operating system (LOS) and the network operating system (NOS).

Servers: Servers are computers that hold shared files, programs and the network operating system. Servers provide access to network resources to all the users of the network and different kinds of servers are present. Examples include file servers, print servers, mail servers, communication servers etc.

Clients: Clients are computers that access and use the network and shared network resources. Client computers are basically the customers (users) of the network, as they request and receive service from the servers.

Shared Data: Shared data are data that file servers provide to clients, such as data files, printer access programs, and e-mail.

Shared Printers and other peripherals: these are hardware resources provided to the users of the network by servers. Resources provided include data files, printers, software, or any other items used by the clients on the network.

Network interface card: Every computer in the network has a special expansion card called network interface card (NIS), which prepares and sends data, receives data, and controls data flow between the computer and the network. While transmitting, NIC passes frames of data on to the physical layer and on the receiver side, the NIC processes bits received from the physical layer and processes the message based on its contents.

Local operating system: A local operating system allows personal computers to access files, print to a local printer, and have and use one or more disk and CD drives that are located on the computer. Examples are MS-DOS, PC-DOS, UNIX, Macintosh, OS/2, Windows 95, 98, XP and Linux.

Network operating system: the NOS is a program that runs on computers and servers that allows the computers to communicate over a network. The NOS provides services to clients such as log-in features, password authentication, printer access, network administration functions and data file sharing.

Network Models

Computer networks can be represented with two basic network models: peer-to-peer client/server and dedicated client/server. The client/server method specifies the way in which two computers can communicate with software over a network.

Peer-to-peer client/server network: Here, all the computers share their resources, such as hard drives, printers and so on with all the other computers on the network. Individual resources like disk drives, CD-ROM drives, and even printers are transformed into shared, collective resources that are accessible from every PC. Unlike client-server networks, where network information is stored on a centralized file server PC and made available to tens, hundreds, or thousands client PCs, the information stored across peer-to-peer networks is uniquely decentralized. Because peer-to-peer PCs have their own hard disk drives that are accessible by all computers, each PC acts as both a client (information requestor) and a server (information provider). The peer-to-peer network is an appropriate choice when there are fewer than 10 users on the network, security is not an issue and all the users are located in the same general area.

The advantages of peer-to-peer over client-server NOSs include:

- No need for a network administrator
- Network is fast/inexpensive to setup & maintain
- Each PC can make backup copies of its data to other PCs for security.
- Easiest type of network to build, peer-to-peer is perfect for both home and office use.

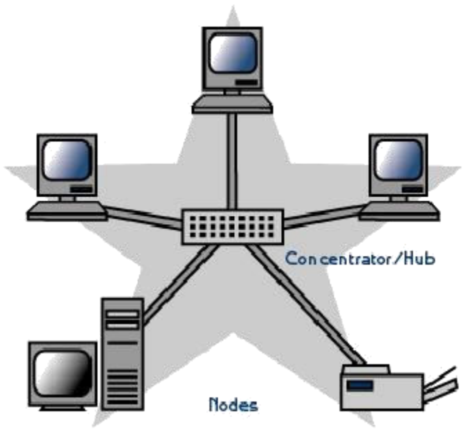
Dedicated client/server network: Here, one computer is designated as server and the rest of the computers are clients. Dedicated Server Architecture can improve the efficiency of client server systems by using one server for each application that exists within an organization. The designated servers store all the networks shared files and applications programs and function only as servers and are not used as a client or workstation. Client computers can access the servers and have shared files transferred to them over the transmission medium. In some client/server networks, client computers submit jobs to one of the servers and once they process the jobs, the results are sent back to the client computer.

In general, the dedicated client/server model is preferable to the peer-to-peer client/server model for general purpose data networks.

Network Topologies

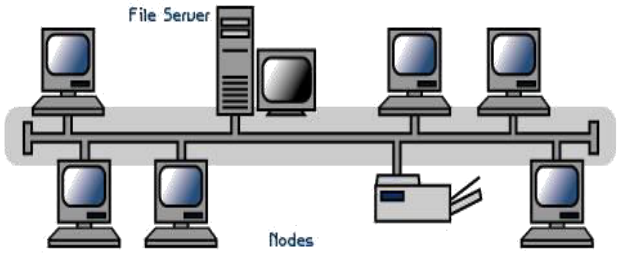
In computer networking, *topology* refers to the layout of connected devices, i.e. how the computers, cables, and other components within a data communications network are interconnected, both physically and logically. The physical topology describes how the network is actually laid out, and the logical topology describes how the data actually flow through the network. Two most basic topologies are point-to-point and multipoint. A point-to-point topology usually connects two mainframe computers for high-speed digital information. A multipoint topology connects three or more stations through a single transmission medium and some examples are *star, bus, ring, mesh and hybrid.*

Star topology: A star topology is designed with each node (file server, workstations, and peripherals) connected directly to a central network hub, switch, or concentrator. Data on a star network passes through the hub, switch, or concentrator before continuing to its destination. The hub, switch, or concentrator manages and controls all functions of the network. It also acts as a repeater for the data flow.



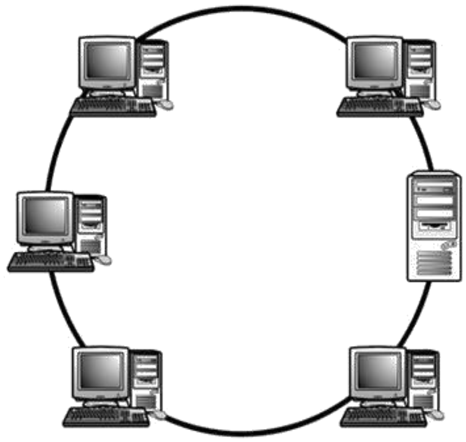
Advantages	Disadvantages
Easily expanded without disruption to the network	Requires more cable
Cable failure affects only a single user	A central connecting device allows for a single point of failure
Easy to troubleshoot and isolate problems	More difficult to implement

Bus topology: Bus networks use a common backbone to connect all devices. A single cable, (the backbone) functions as a shared communication medium that devices attach or tap into with an interface connector. A device wanting to communicate with another device on the network sends a broadcast message onto the wire that all other devices see, but only the intended recipient actually accepts and processes the message. The bus topology is the simplest and most common method of interconnecting computers. The two ends of the transmission line never touch to form a complete loop. A bus topology is also known as multidrop or linear bus or a horizontal bus.



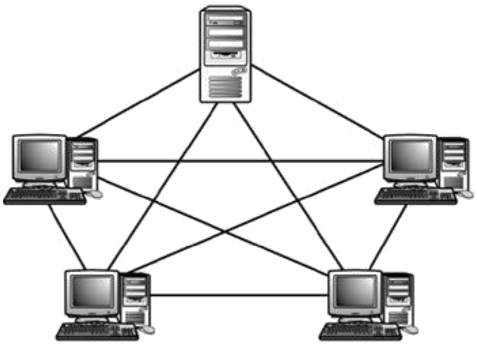
Advantages	Disadvantages
Cheap and easy to implement	Network disruption when computers are added or removed
Require less cable	A break in the cable will prevent all systems from accessing the network.
Does not use any specialized network equipment.	Difficult to troubleshoot.

Ring topology: In a ring network (sometimes called a loop), every device has exactly two neighbours for communication purposes. All messages travel through a ring in the same direction (either "clockwise" or "counter clockwise"). All the stations are interconnected in tandem (series) to form a closed loop or circle. Transmissions are unidirectional and must propagate through all the stations in the loop. Each computer acts like a repeater and the ring topology is similar to bus or star topologies.



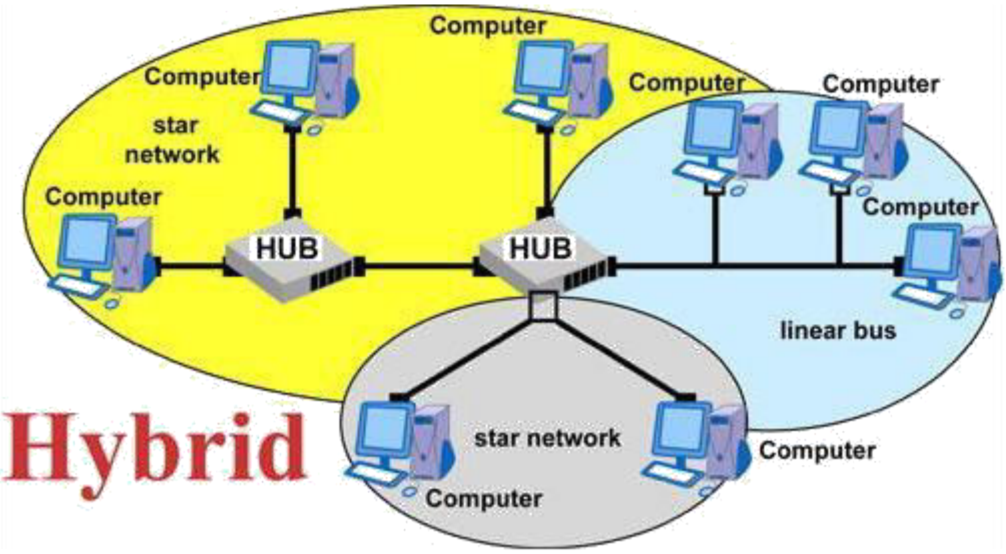
Advantages	Disadvantages
Cable faults are easily located, making troubleshooting easier	Expansion to the network can cause network disruption
Ring networks are moderately easy to install	A single break in the cable can disrupt the entire network.

Mesh topology: The *mesh* topology incorporates a unique network design in which each computer on the network connects to every other, creating a point-to-point connection between every device on the network. Unlike each of the previous topologies, messages sent on a mesh network can take any of several possible paths from source to destination. A mesh network in which every device connects to every other is called a full mesh. A disadvantage is that, a mesh network with *n* nodes must have *n(n-1)/2* links and each node must have *n-1* I/O ports (links).



Advantages	Disadvantages
Provides redundant paths between devices	Requires more cable than the other LAN topologies
The network can be expanded without disruption to current uses	Complicated implementation

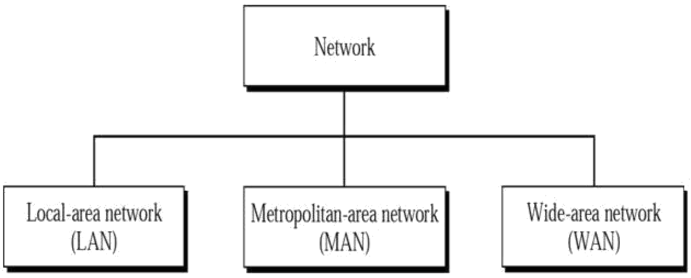
Hybrid topology: This topology (sometimes called mixed topology) is simply combining two or more of the traditional topologies to form a larger, more complex topology. Main aim is being able to share the advantages of different topologies.

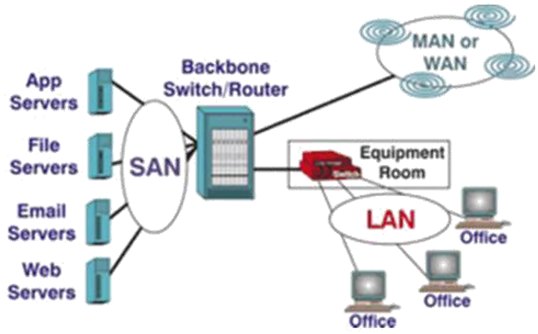


Network Classifications

One way to categorize the different types of computer network designs is by their scope or scale. Common examples of area network types are:

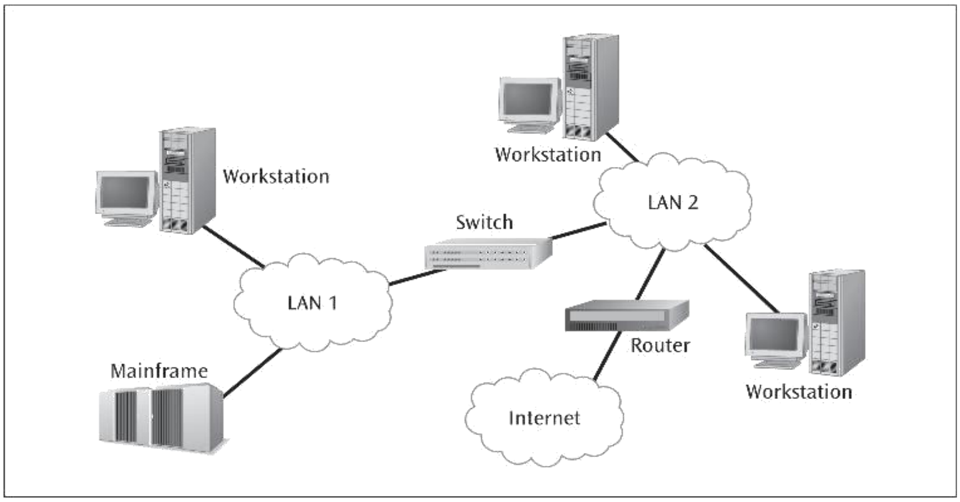
- LAN - Local Area Network
- WLAN - Wireless Local Area Network
- WAN - Wide Area Network
- MAN - Metropolitan Area Network
- SAN - Storage Area Network, System Area Network, Server Area Network, or sometimes Small Area Network
- CAN - Campus Area Network, Controller Area Network, or sometimes Cluster Area Network
- PAN - Personal Area Network
- DAN - Desk Area Network





Local area network: A local area network (LAN) is a network that connects computers and devices in a limited geographical area such as home, school, computer laboratory, office building, or closely positioned group of buildings. LANs use a network operating system to provide two-way communications at bit rates in the range of 10 Mbps to 100 Mbps. In addition to operating in a limited space, LANs are also typically owned, controlled, and managed by a single person or organization. They also tend to use certain connectivity technologies, primarily Ethernet and Token Ring.

Figure 7-1
A local area network interconnecting another local area network, the Internet, and a mainframe computer



Advantages of LAN:

- ✓ Share resources efficiently
- ✓ Individual workstation might survive network failure if it doesn't rely upon others
- ✓ Component evolution independent of system evolution
- ✓ Support heterogeneous hardware/software
- ✓ Access to other LANs and WANs
- ✓ High transfer rates with low error rates

Metropolitan area network: A MAN is optimized for a larger geographical area than a LAN, ranging from several blocks of buildings to entire cities. Its geographic scope falls between a WAN and LAN. A MAN might be a single network like the cable television network or it usually interconnects a number of local area networks (LANs) using a high-capacity backbone technology, such as fiber-optical links, and provides up-link services to wide area networks and the Internet. MANs typically operate at speeds of 1.5 Mbps to 10 Mbps and range from five miles to a few hundred miles in length. Examples of MANs are FDDI (fiber distributed data interface) and ATM (asynchronous transfer mode).

~~**Wide area network**~~: Wide area networks are the oldest type of data communications network that provide relatively slow-speed, long-distance transmission of data, voice and video information over relatively large and widely dispersed geographical areas, such as country or entire continent. WANs interconnect routers in different locations. A WAN differs from a LAN in several important ways. Most WANs (like the Internet) are not owned by any one organization but rather exist under collective or distributed ownership and management. WANs tend to use technology like ATM, Frame Relay and X.25 for connectivity over the longer distances.

~~**Global area network**~~: A GAN provides connections between countries around the entire globe. Internet is a good example and is essentially a network comprised of other networks that interconnect virtually every country in the world. GANs operate from 1.5 Mbps to 100 Gbps and cover thousands of miles.

Campus Area Network: - a network spanning multiple LANs but smaller than a MAN, such as on a university or local business campus.

Storage Area Network: - connects servers to data storage devices through a technology like Fibre Channel.

System Area Network: - Links high-performance computers with high-speed connections in a cluster configuration. Also known as Cluster Area Network.

Building backbone: - It is a network connection that normally carries traffic between departmental LANs within a single company. It consists of a switch or router to provide connectivity to other networks such as campus backbones, enterprise backbones, MANs, WANs etc

Camus backbone: - **It** is a network connection used to carry traffic to and from LANs located in various buildings on campus. It normally uses optical fiber cables for the transmission media between buildings and operates at relatively high transmission rates.

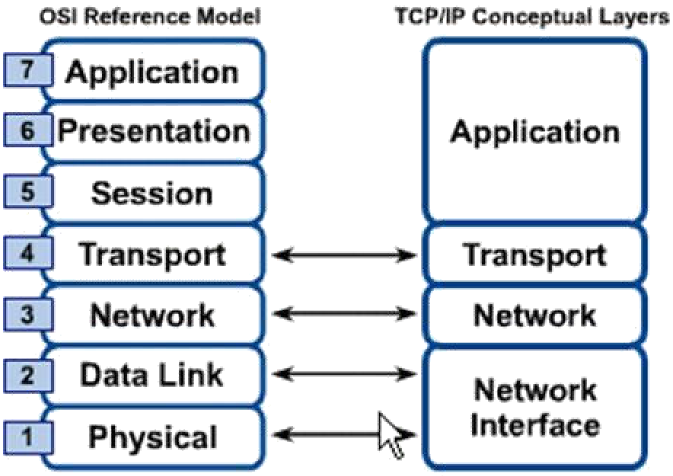
Enterprise networks: - It includes some or all of the above networks and components connected in a cohesive and manageable fashion.

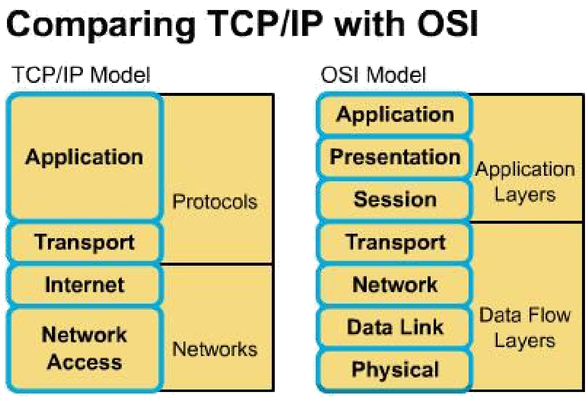
Alternate Protocol Suites

The protocols other than OSI that are in wide spread used are TCP?IP and the Cisco three-layer hierarchical model.

TCP/IP Protocol Suite

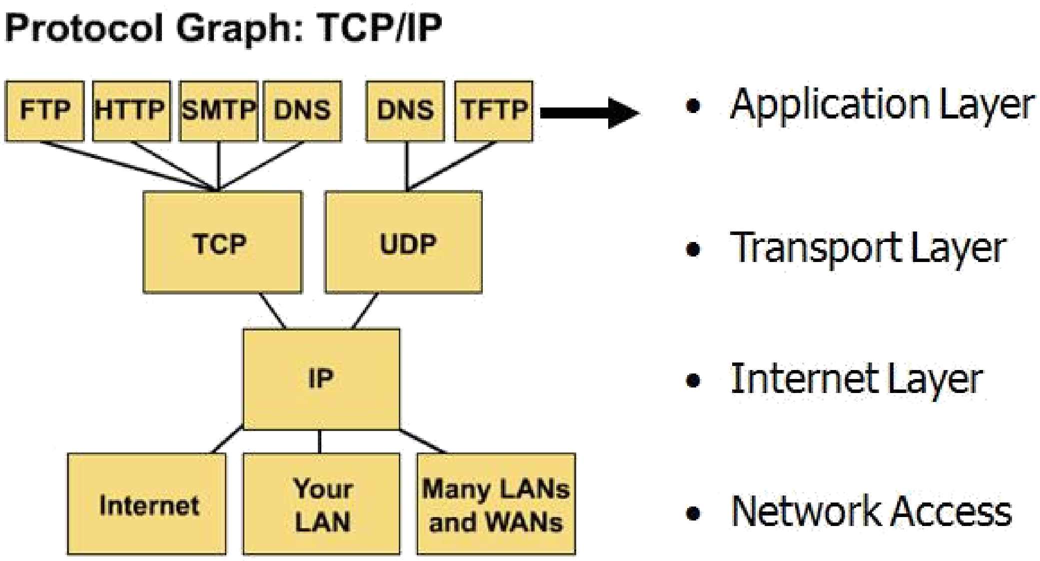
The U.S. Department of Defense (*DoD*) created the TCP/IP reference model because it wanted a network that could survive any conditions, even a nuclear war. Transmission Control Protocol/Internet Protocol (TCP/IP) {commonly known as internet suite} model is a set of communication protocols that allow communication across multiple diverse networks. TCP/IP is a hierarchical protocol comprised of either three or four layers. The three-layer version of TCP/IP contains the network, transport and application layers. Four layer version specifies the host to network layer.





The designers of TCP/IP felt that the higher level protocols should include the *session* and *presentation* layer details. They simply created an **application** layer that handles high-level protocols, issues of representation, encoding, and dialog control. The TCP/IP combines all application-related issues into one layer, and assures this data is properly packaged for the next layer.

The TCP/IP **transport layer** deals with the quality-of-service issues of reliability, flow control, and error correction. One of its protocols, the transmission control protocol (TCP), provides excellent and flexible ways to create reliable, well-flowing, low-error network communications. TCP is a *connection-oriented protocol*. The other protocol is User Datagram Protocol (UDP) which is a connection less protocol.



Common TCP/IP Protocols

The purpose of the ***Internet layer*** is to send source packets from any network on the internetwork and have them arrive at the destination independent of the path and networks they took to get there. The specific protocol that governs this layer is called the **Internet protocol** (IP). *Best path determination* and *packet switching* occur at this layer.

The **network access layer** also called the *host-to-network* layer is concerned with all of the issues of physically delivering data packets using frames or cells.

Differences between OSI and TCP/IP

- TCP/IP combines the presentation and session layer issues into its application layer
- TCP/IP combines the OSI data link and physical layers into one layer
- TCP/IP appears simpler because it has fewer layers
- TCP/IP protocols are the standards around which the Internet developed, so the TCP/IP model gains credibility just because of its protocols. In contrast, typically networks aren't built on the OSI protocol, even though the OSI model is used as a guide.

Cisco Three Layer Model

Cisco has defined a hierarchical model known as the hierarchical internetworking model. This model simplifies the task of building a reliable, scalable, and less expensive hierarchical internetwork because rather than focusing on packet construction; it focuses on the three functional areas, or layers, of your network.

Core layer: This layer is considered the backbone of the network and includes the high-end switches and high-speed cables such as fiber cables. This layer of the network does not route traffic at the LAN. In addition, no packet manipulation is done by devices in this layer. Rather, this layer is concerned with speed and ensures reliable delivery of packets.

Distribution layer: This layer includes LAN-based routers and layer 3 switches. This layer ensures that packets are properly routed between subnets and VLANs in your enterprise. This layer is also called the Workgroup layer. It also provides policy-based network connectivity, including:

- Packet filtering (firewalling): Processes packets and regulates the transmission of packets based on its source and destination information to create network borders.
- QoS: The router or layer 3 switches can read packets and prioritize delivery, based on policies set.
- Access Layer Aggregation Point: The layer serves the aggregation point for the desktop layer switches.
- Control Broadcast and Multicast: The layer serves as the boundary for broadcast and multicast domains.
- Application Gateways: The layer allows you to create protocol gateways to and from different network architectures.
- The distribution layer also performs queuing and provides packet manipulation of the network traffic.

Access layer: This layer includes hubs and switches. This layer is also called the desktop layer because it focuses on connecting client nodes, such as workstations to the network. This layer ensures that packets are delivered to end user computers. At the access layer, you can:

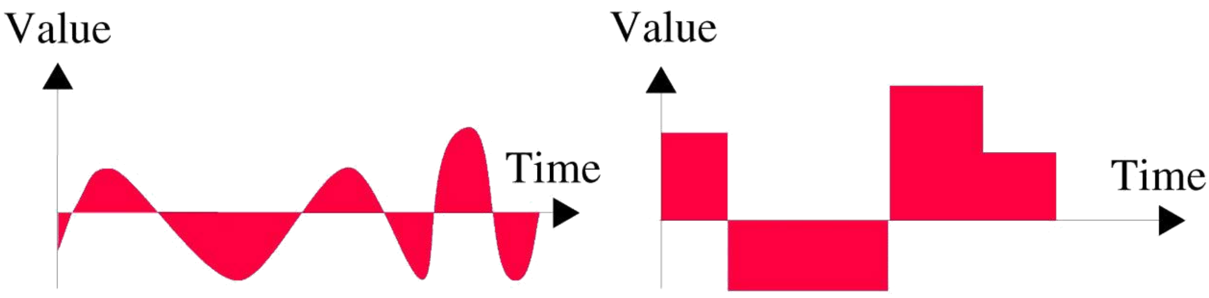
- Enable MAC address filtering: It is possible to program a switch to allow only certain systems to access the connected LANs.
- Create separate collision domains: A switch can create separate collision domains for each connected node to improve performance.
- Share bandwidth: You can allow the same network connection to handle all data.
- Handle switch bandwidth: You can move data from one network to another to perform load balancing.

The benefits of the Cisco hierarchical model includes:

- High Performance: You can design high performance networks, where only certain layers are susceptible to congestion.
- Efficient management & troubleshooting: Allows you to efficiently organize network management and isolate causes of network trouble.
- Policy creation: You can easily create policies and specify filters and rules.
- Scalability: You can grow the network easily by dividing your network into functional areas.
- Behavior prediction: When planning or managing a network, the model allows you determine what will happen to the network when new stresses are placed on it.

Signals, Noise, Modulation and Demodulation

Computers transmit data using digital signals, sequences of specified voltage levels. Computers sometimes communicate over telephone lines using analog signals, which are formed by continuously varying voltage levels. Electrical signals can be in analog or digital form. With analog signals, the amplitude changes continuously with respect to time with no breaks or discontinuities. A sine wave is the most basic analog signal.



a. Analog signal b. Digital signal

Digital signals are described as discrete; their amplitude maintains a constant level for a prescribed period of time and then it changes to another level. If only two levels are possible, it is called a binary signal. All binary signals are digital, but all digital signals are not necessarily binary. Converting information signals to a different form is called ***modulation*** and the reverse process is called ***demodulation***. The modulating signal is the information and the signal being modulated is the ***carrier***.

Two basic types of electronic communications systems are analog and digital. An analog digital communications system is a communications system in which energy is transmitted and received in analog form and are also propagated through the system in analog form. Digital communications covers a broad range of communications techniques including digital transmission and digital modulation.

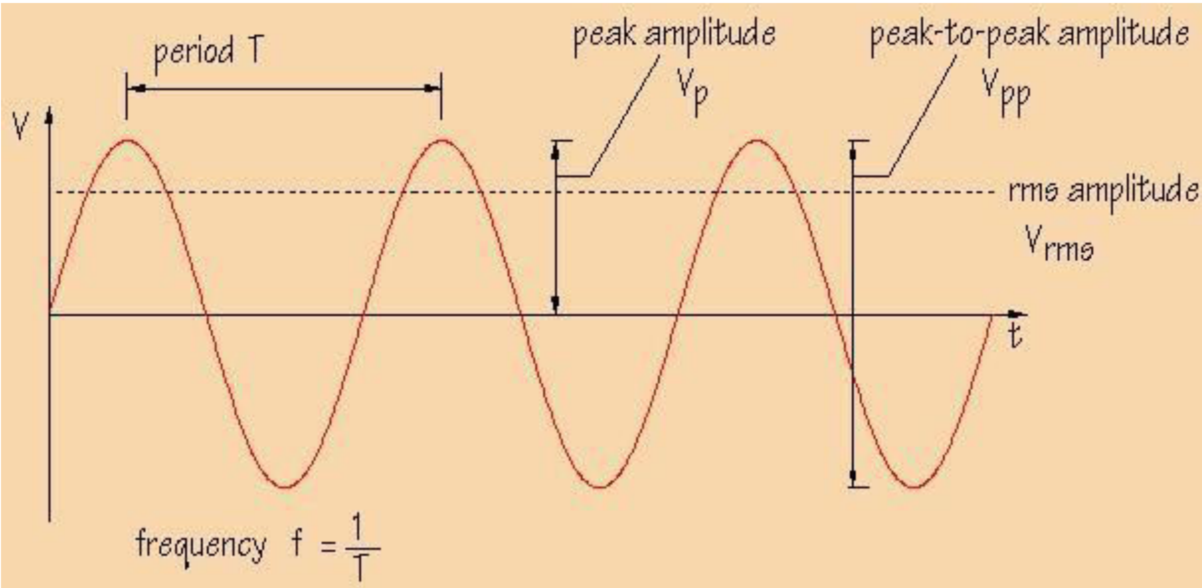
Signal Analysis

Mathematical signal analysis is used to analyze and predict the performance of the circuit on the basis of the voltage distribution and frequency composition of the information signal.

Amplitude, Frequency and Phase

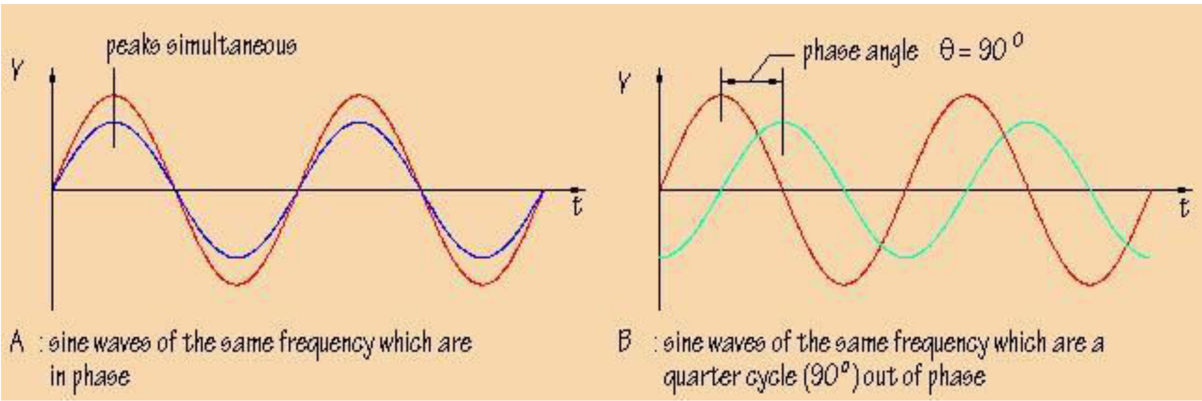
A ***cycle*** is one complete variation in the signal, and the ***period*** is the time the waveform takes to complete on cycle. One cycle constitutes 360 degrees (or 2π radians). Sine waves can be described in terms of three parameters: ***amplitude, frequency and phase.***

Amplitude (A): It is analogous to magnitude or displacement. The amplitude of a signal is the magnitude of the signal at any point on the waveform. The amplitude of electrical signal is generally measured in voltage. The maximum voltage of a signal in respect to its average value is called its peak amplitude or peak voltage.

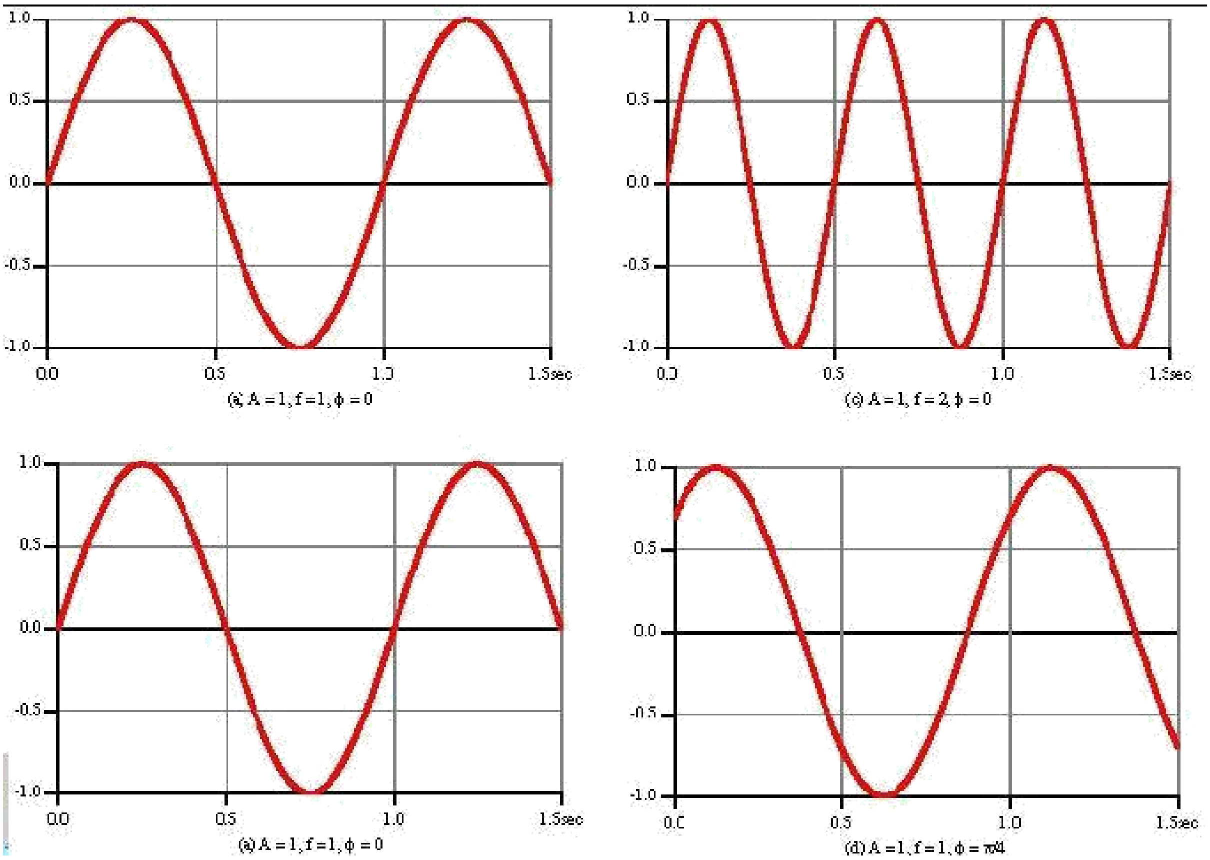


Frequency (f): The time of one cycle of a waveform is its period, which is measured in seconds. Frequency is the number of cycles completed per second. The measurement unit for frequency is the **hertz**, **Hz**. 1 Hz = 1 cycle per second. The frequency of the signal can be calculated from**T=1/*f***

Phase (Ø): The phase of the signal is measured in degrees or radians with respect to a reference point. A phase shift of 180 degrees corresponds to a shift of half a cycle.



A phase shift of 360 degrees corresponds to a shift of one complete cycle. If two sine waves have the same frequency and occur at the same time, they are said to be **in phase**, or they are said to out of phase. The difference in phase can be measured in degrees, and is called the **phase angle**, θ



Varying Sine wave with respect to frequency and phase

Periodic Signals

A signal is periodic if it completes a pattern within a measurable time and is characterized by amplitude, frequency and phase. Mathematically, a single frequency voltage waveform is

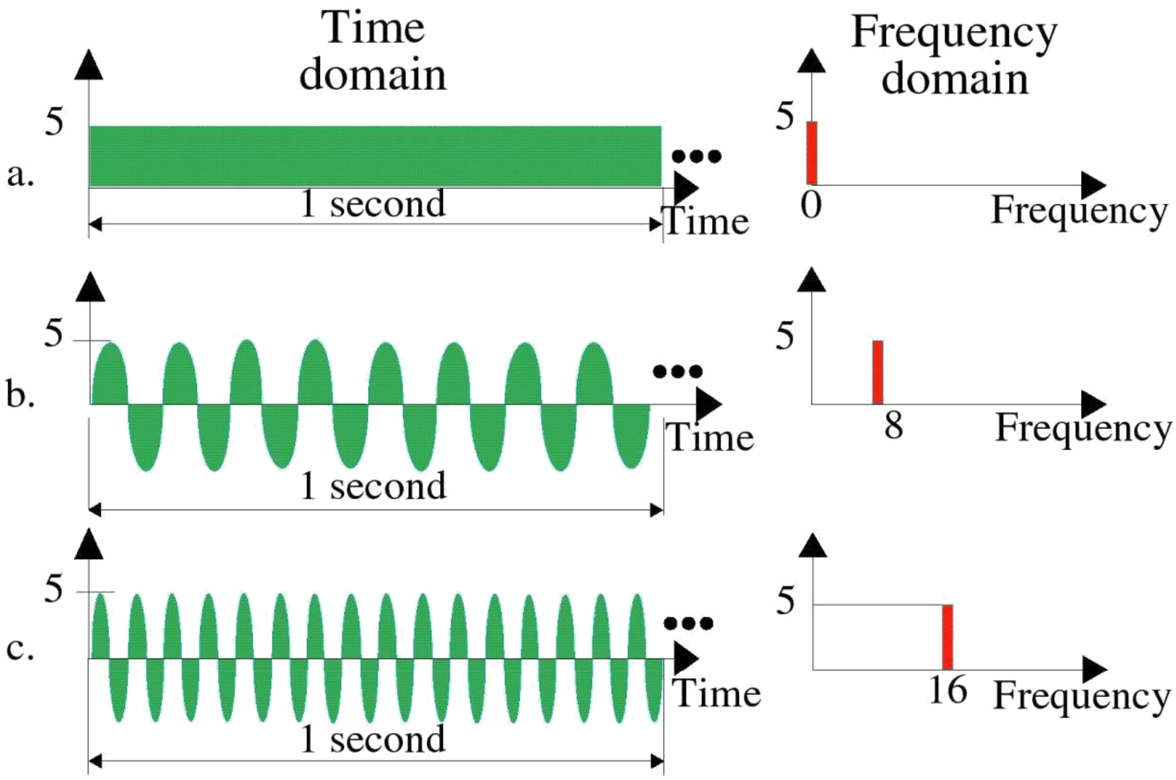
$v(t) = V \sin(2\pi f t + \theta)$, where

- *v(t)* is time-varying voltage sine wave
- V is peak amplitude in volts
- f is frequency in hertz
- t is time in seconds
- θ is phase in degrees or radians

It is called a periodic wave because, it repeats at a uniform rate. A series of sine, cosine or square waves constitute an example of periodic waves, which can be analyzed in either the time domain or the frequency domain.

Time domain: Time domain is a term used to describe the analysis of mathematical functions, or physical signals, with respect to time. In the time domain, the signal or function's value is known for all real numbers, for the case of continuous time, or at various separate instants in the case of discrete time. An oscilloscope is a time-domain tool commonly used to visualize real-world signals in the time domain. A time domain graph shows how a signal changes over time.

Frequency Domain: frequency domain is a term used to describe the analysis of mathematical functions or signals with respect to frequency, rather than time. A spectrum analyser is a frequency-domain instrument which displays amplitude-versus frequency plot (called a frequency spectrum). The horizontal axis represents frequency and the vertical axis amplitude showing a vertical deflection for each frequency present in the waveform, which is proportional to the amplitude of the frequency it represents.



Examples for time domain and frequency domain

Complex Signals

Any repetitive waveform that is comprised of more than one harmonically related sine or cosine wave is called a nonsinusoidal, complex wave. Fourier series is used to analyze the complex periodic waves.

Fourier series: The Fourier series is used in signal analysis to represent the sinusoidal components of nonsinusoidal periodic waveforms. A **Fourier series** decomposes a periodic function or periodic signal into a sum of simple oscillating functions, namely sines and cosines. It can be expressed as:

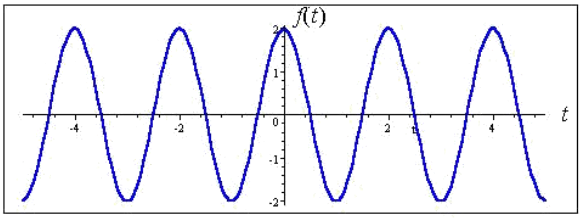
$$f(t) = A_0 + A_1 \cos\alpha + A_2 \cos 2\alpha + A_3 \cos 3\alpha + \ldots\ldots A_n \cos n\alpha +$$
$$A_0 + B_1 \sin\beta + B_2 \sin 2\beta + B_3 \sin 3\beta + \ldots\ldots\ldots B_n \sin n\beta$$

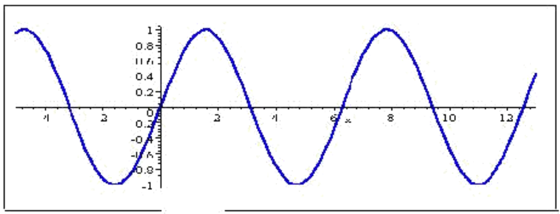
Where $\alpha=\beta$

Any periodic waveform is comprised of an average dc component and a series of harmonically related sines or cosine waves. A harmonic is an integral multiple of the fundamental frequency. Fundamental frequency is the first harmonic and equal to the frequency (repetition rate) of the waveform. Second multiple is called second harmonic, third multiple is called third harmonic and so forth.

Wave symmetry: It describes the symmetry of a waveform in the time domain, i.e., its relative position with respect to the horizontal (time) and vertical (amplitude) axes.

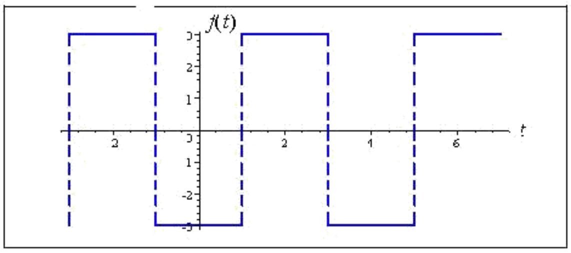
Even symmetry: If a periodic voltage waveform is symmetric about the vertical axis, it is said to have axes, or mirror, symmetry and is called an even function. For all even functions, the **β** coefficients are zero. Even function satisfy the condition *f(t) = f(-t)*

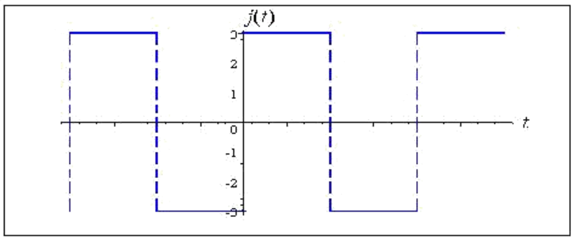




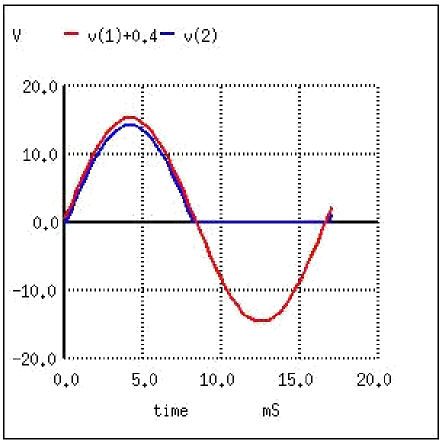
Examples of Even waves: sine wave and square wave

Examples of Odd waves: sine wave and square wave





Odd symmetry: If a periodic voltage waveform is symmetric about a line midway between the vertical axis and the negative horizontal axis and passing through the coordinate origin, it is said to have to point or skew, symmetry and is called an odd function. For all odd functions, the α coefficients are zero. Odd function satisfies *f(t) = -f(-t)*



Half-wave symmetry: If a periodic voltage waveform is such that the waveform for the first half cycle repeats itself except with the opposite sign for the second half cycle, it is called to have half-wave symmetry. Half-wave symmetry implies that the second half of the wave is exactly opposite to the first half. A function with half-wave symmetry does not have to be even or odd, as this property requires only that the shifted signal is opposite. Half-wave functions satisfy the condition *f(t) = -f(T+t)/2*

Frequency Spectrum and Bandwidth

The frequency spectrum of a waveform consists of all the frequencies contained in the waveform and their respective amplitudes plotted in the frequency domain.

Bandwidth of an information signal is simply the difference between the highest and lowest frequencies contained in the information and the bandwidth of a communication channel is the difference between the highest and lowest frequencies that the channel will allow to pass through it.

Electrical Noise and Signal-To-Noise Ratio

Noise is any disturbance or distortion that comes in the process of communication. Electrical noise is defined as any undesirable electrical energy that falls within the passband of the signal. A noise signal consists of a mixture of frequencies with random amplitudes. Noise can originate in various ways. The most prevalent and most interfering to data communication signals are *man-made noise, thermal noise, correlated noise, and impulse noise.*

Man-made noise: It is the kind of noise produced by mankind. The main sources are spark-producing mechanisms like commutators in electric motors, automobile ignition systems, ac power-generating and switching equipment, and fluorescent lights. It is impulsive in nature and contains a wide range of frequencies propagated in the free space like the radio waves. Man-made noise is most intense in more densely populated areas and sometimes is referred to as *industrial noise.*

Thermal noise: This is the noise generated by thermal agitation of electrons in a conductor. It is also referred to as white noise because of its uniform distribution across the entire electromagnetic frequency spectrum. Noise power density is the thermal noise power present in a 1-Hz bandwidth and is given by $N_O = KT$.

Thermal noise is independent of frequency and thus thermal noise present in any bandwidth is

$$N = KTB$$

where **N** is thermal noise power in watts, **K** is Boltzmann's constant in joules per Kelvin, ***T*** is the conductor temperature in kelvin ($0K = -273^{O}C$), and **B** is the bandwidth in hertz. Noise power is often measured in dBm. From the equation above, noise power in a resistor at room temperature, in dBm, is: $\mathbf{N_{dBm} = -174\ dBm + 10 \log B}$

Correlated noise: this noise is correlated to the signal and cannot be present in a circuit unless there is a signal. Correlated noise is produced by nonlinear amplification and includes harmonic distortion and intermodulation distortion. Harmonic distortion occurs when unwanted harmonics of a signal are produced through nonlinear amplification and is also

called amplitude distortion. Intermodulation distortion is the generation of unwanted sum and difference frequencies produced when two or more signals are amplified in a nonlinear device.

Impulse noise: This noise is characterized by high-amplitude peaks of short duration in the total noise spectrum. It consists of sudden bursts of irregularly shaped pulses that generally last between a few microseconds and several milliseconds, depending on their amplitude and origin. In case of voice communications, impulse noise is very annoying as it generates a sharp popping or crackling sound where as it is devastating in data circuits.

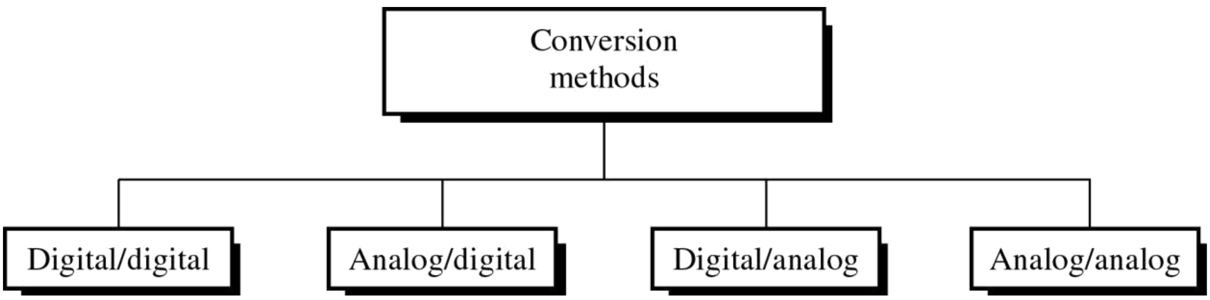
Signal-to-noise power ratio: **Signal-to-noise ratio** (often abbreviated **SNR** or **S/N**) is defined as the ratio of signal power to the noise power corrupting the signal. A ratio higher than 1:1 indicates more signal than noise. Signal-to-noise ratio is defined as the power ratio between signal (meaningful information) and the background noise (unwanted signal)

$$\mathrm{SNR} = \frac{P_{\mathrm{signal}}}{P_{\mathrm{noise}}},$$

where *P* is average power in watts. The ratio often expressed in decibels as $S/N_{(dBm)} = 10 \log(P_S/P_N)$

Analog Modulation Systems

A sine wave has three main components: amplitude, frequency and phase and can be expressed as $v(t) = V \sin(2\pi f t + \theta)$. If the information signal is analog and the amplitude (V) of the carrier is varied proportional to the informational signal, *amplitude modulation* (AM) is produced. If the frequency (f) is varied proportional to the information signal, *frequency modulation* (FM) is produced and if the phase (θ) is varied proportional to the information signal, *phase modulation* (PM) is produced. Frequency and phase modulation are similar and often combined and are simply called *angle modulation*.

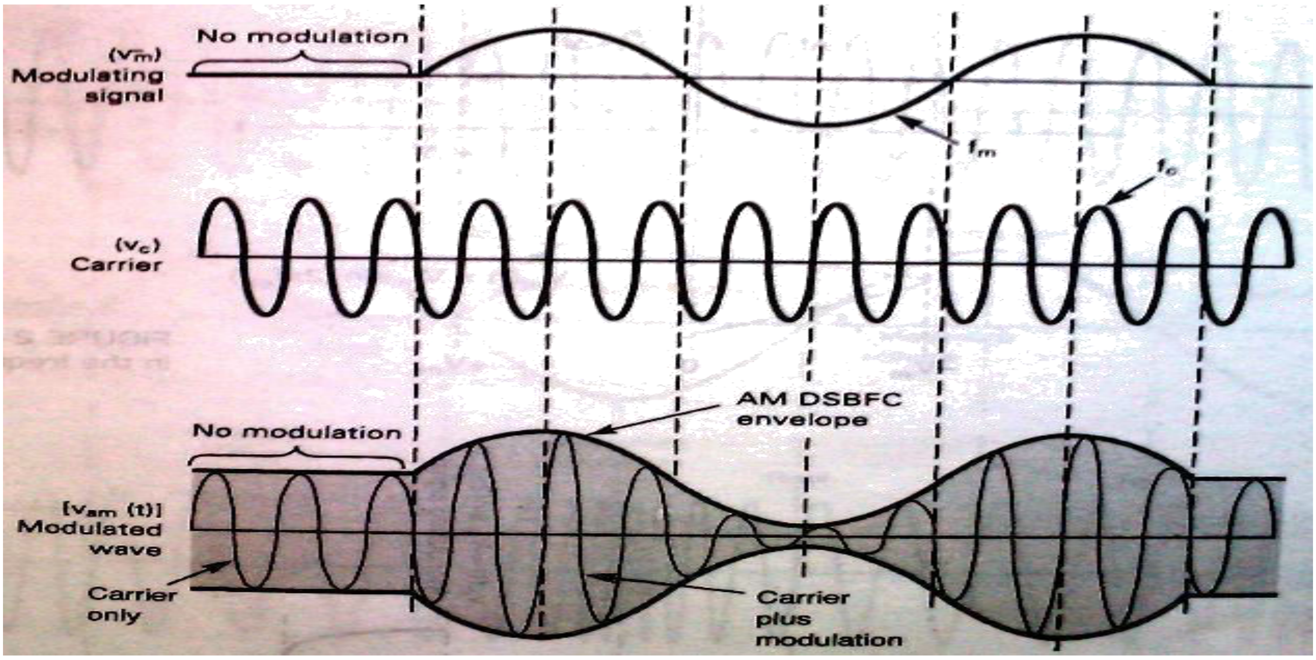


The process of impressing relatively low-frequency information signals onto a high-frequency carrier signal is called *modulation* and the reverse process is called *demodulation*.

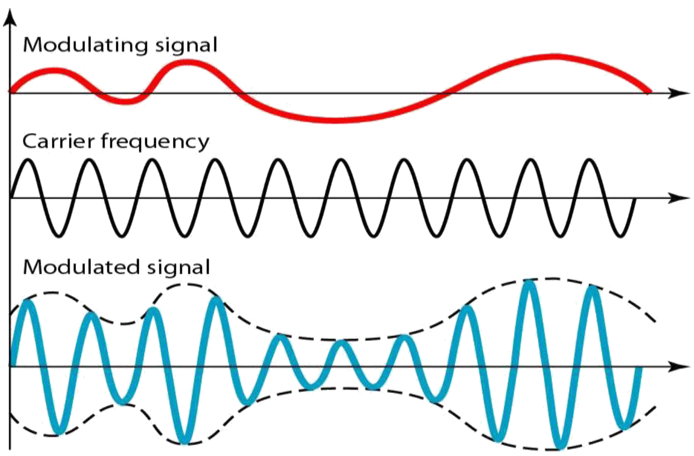
Analog modulation is used for the transmission of conventional analog signals, such as voice, music, and video and not particularly useful for data communication systems.

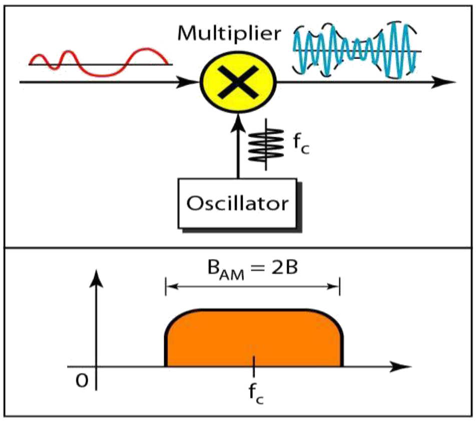
Amplitude Modulation

Amplitude modulation is the process of changing the amplitude of a relatively high frequency carrier signal in proportion to the instantaneous value of the modulating signal (information). AM modulators are two-input devices, one of them is a single, relatively high frequency carrier signal of constant amplitude and the second is the relatively low-frequency information signal. The following figure shows generation of AM waveform when a single-frequency modulating signal acts on a high frequency carrier signal.



AM generation





Advantages of AM are simple to implement, needs a circuit with very few components and inexpensive. The disadvantages include inefficient power usage and use of bandwidth and also prone to noise. The total bandwidth required for AM can be determined from the bandwidth of the audio signal: $B_{AM} = 2B$

Angle Modulation

Angle modulation results whenever the phase angle of a sinusoidal signal is varied with respect to time and includes both FM and PM. Whenever the frequency of a carrier signal is varied, the phase is also varied and vice versa. If the frequency of the carrier is varied directly in accordance with the information signal, FM results, whereas if the phase is varied directly, PM results.

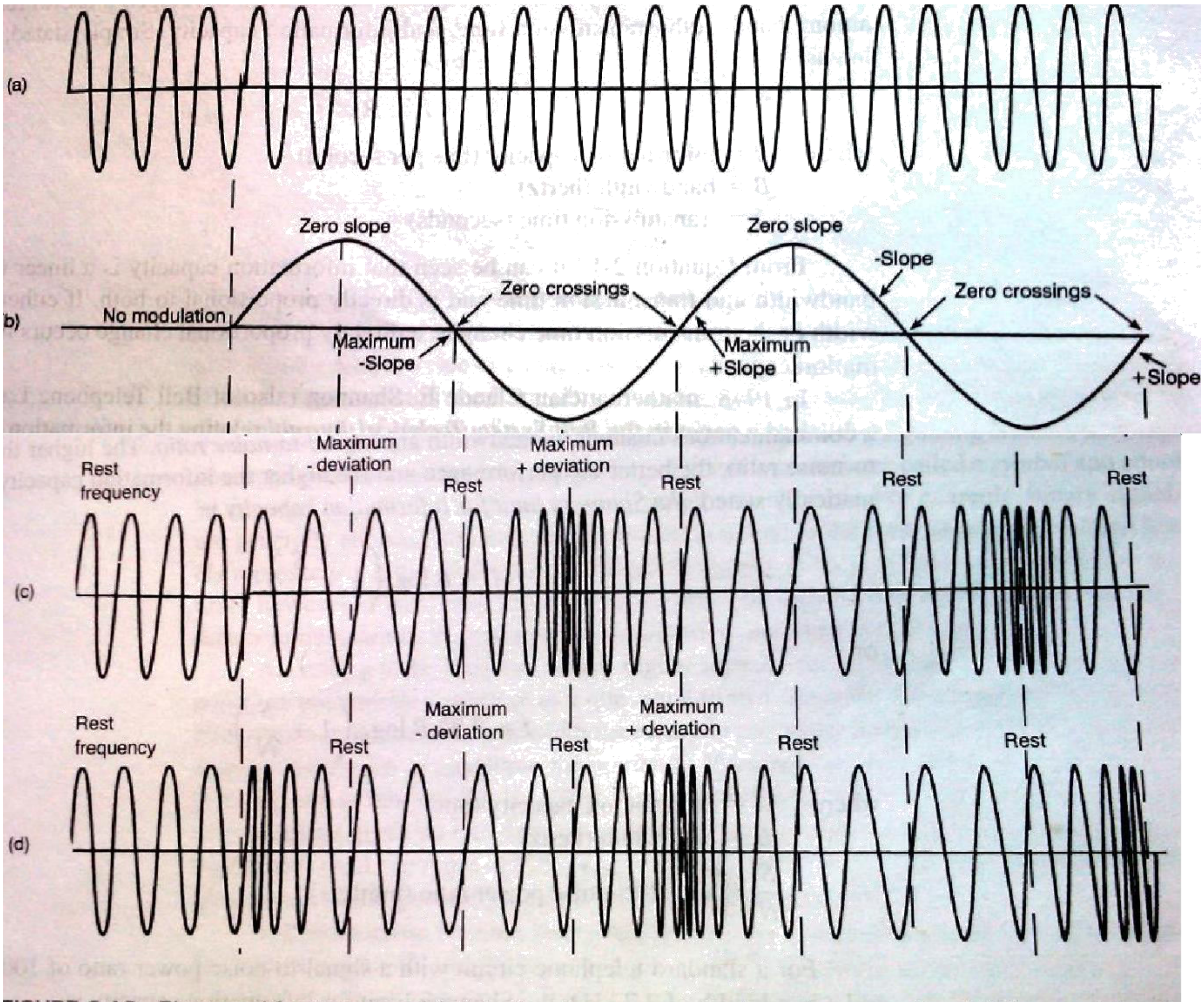
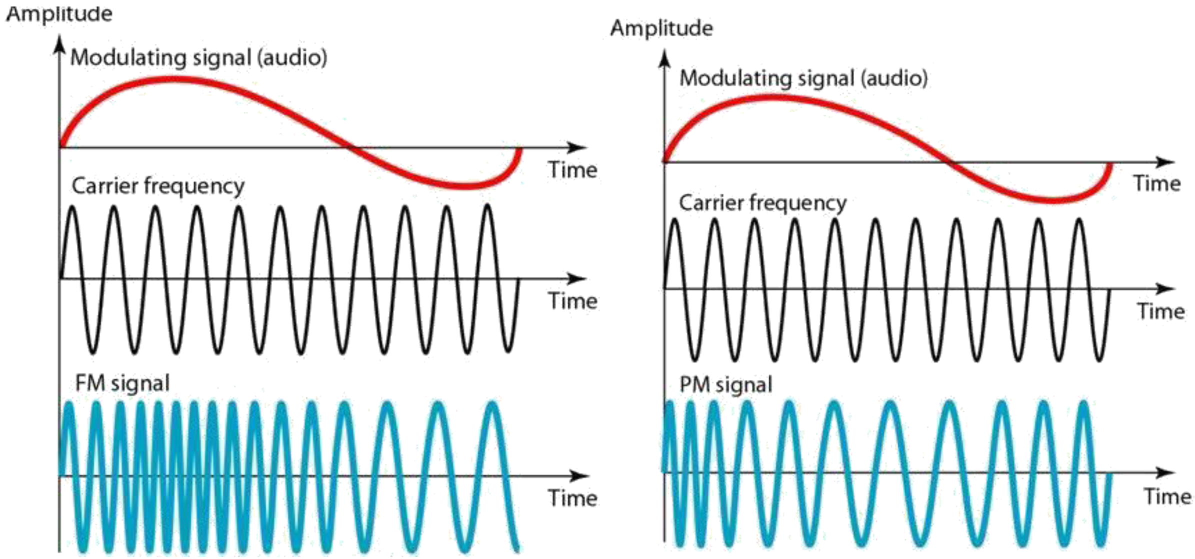


FIGURE 2-16 Phase and frequency modulation of a sine-wave carrier by a sine-wave signal: (a) unmodulated carrier; (b) modulating signal; (c) frequency-modulated wave; (d) phase-modulated wave

The above figure shows the FM and PM of a sinusoidal carrier by a single-frequency modulating signal. Both FM and PM waveforms are identical except for their time relationship (phase). With FM, the maximum frequency deviation occurs during the maximum positive and negative peaks of the modulating signal. With FM, the maximum frequency deviation occurs during the zero crossings in the modulating signal.



Comparison of FM and PM

An important feature of FM and PM is that they can provide much better protection to the message against channel noise when compared to AM. Also because of their constant amplitude nature, they can withstand nonlinear distortion and amplitude fading.

Information Capacity, Bits, Bit Rate, Baud, and M-ARY Encoding

Information capacity is a measure of how much information can be propagated through a communication system and a function of bandwidth and transmission time. It represents the number of independent symbols that can be carried through a system in a given unit of time. The most basic digital symbol used to represent information is the ***binary digit, or bit***. ***Bit rate*** is simply the number of bits transmitted during 1 second and is expressed as *bits per second (bps).*

R.Hartley developed a useful relationship among bandwidth, transmission time and information capacity called **Hartley's law** given by:

$$I \propto B \times t$$

Where, *I is the information capacity in bps, B is bandwidth in hertz and t is transmission time in sec's*

Relation between information capacity of a communication channel to a bandwidth and signal-to-noise ratio is given by Claude E. Shannon. The higher the signal-to-noise ratio, the better the performance and also information capacity is higher. The **Shannon limit of information capacity** is

$$I = B \log_2 (1 + S/N) \quad \text{or} \quad I = 3.32\, B \log_{10} (1 + S/N)$$

Where I is information capacity in bps, B is bandwidth in hertz and S/N is signal to noise ratio.

M-ary Encoding

M-ary is a term derived from the word binary. M simply represents a digit that corresponds to the number of conditions, levels, or combinations possible for a given number of binary variables. For example, a digital signal with four possible conditions is an M-ary system where M= 4 and if there are eight possible conditions, then M= 8. The number of bits necessary to produce a given number of conditions is expressed mathematically as:

$\mathbf{N = \log_2 M}$ or it can be written as $\mathbf{M = 2^N}$**,** where N is no of bits necessary and M is number of conditions, levels or combinations possible with N bits. From the equation, it can be said that if there is one bit, only 2^1 or two conditions are possible. For two bits 2^2 or four conditions are possible.

Baud and Minimum Bandwidth

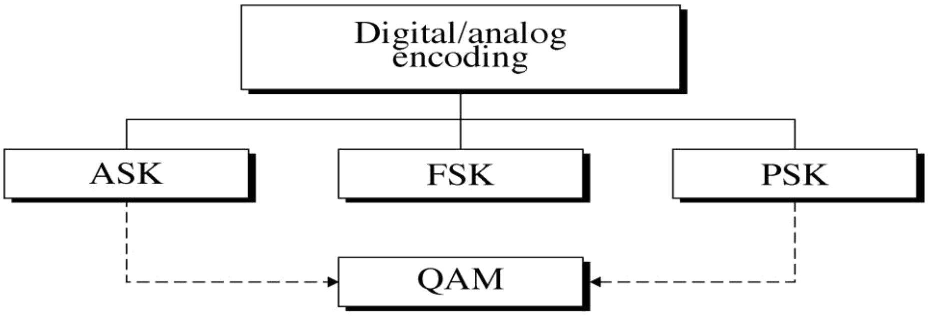
Baud, like bit rate is a rate of change. Baud refers to the rate of change of the signal on the transmission medium after encoding and modulation have occurred. Baud is the reciprocal of the time of one output signalling element, and a signalling element may represent several information bits. Baud is also transmitted one at a time and a baud may represent more than one information bit. So, the baud of the data communications system may be considerably less than the bit rate.

According to H.Nyquist, binary digital signals can be propagated through an ideal noiseless medium at a rate equal to twice the bandwidth of the medium. The minimum theoretical bandwidth necessary to propagate a signal is called the minimum Nyquist bandwidth or sometimes the Nyquist bandwidth. Using multilevel signalling, the Nyquist formulation for channel capacity is $\boldsymbol{f_b = B \log_2 M}$ where, f_b is channel capacity in bps, B is minimum Nyquist bandwidth in hertz and M is no of discrete signal or voltage levels. If N is substituted, we get

$\mathbf{B = baud = } \boldsymbol{f_b/N}$**,** where N is number of bits encoded into each signalling element.

Digital Modulation

Digital modulation is the transmission of digitally modulated analog signals between two or more points in a communications system. Analog and Digital modulation systems use analog carriers to transport information through the system, but digital modulation uses digital modulating (information) signal. Analog systems use analog signal only. In, $\boldsymbol{v(t) = V \sin(2\pi ft + \theta)}$, if the information signal is digital and amplitude (V) of the carrier is varied proportional to the information signal, a digitally modulated signal called amplitude-shift keying (ASK) is produced. If the frequency (f) is varied proportional to the information signal, frequency-shift keying (FSK) is produced and if the phase is varied proportional to the information signal, phase-shift keying (PSK) is produced. If both amplitude and phase are varied proportional to the information signal, quadrature amplitude modulation (QAM) results.



Digital modulation is ideally suited to a multitude of communications applications including both cable and wireless systems. Applications include relatively low-speed voice-band data communications systems, high-speed data transmission systems, digital satellite communication systems and personal communication systems (PCS).

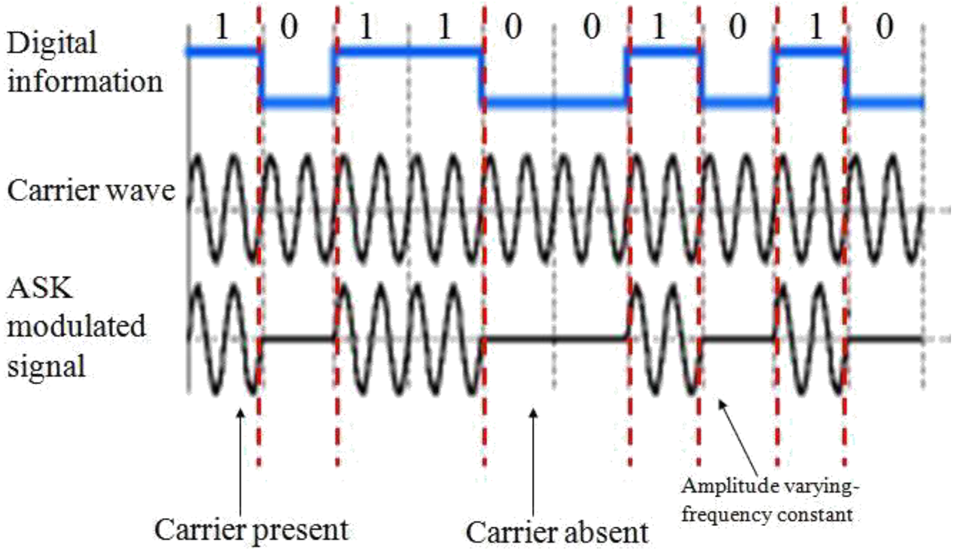
Modulation format	Application
MSK, GMSK	GSM, CDPD
BPSK	Deep space telemetry, cable modems
QPSK, $^{\pi}/_{4}$ DQPSK	Satellite, CDMA, NADC, TETRA, PHS, PDC, LMDS, DVB-S, cable (return path), cable modems, TFTS
OQPSK	CDMA, satellite
FSK, GFSK	DECT, paging, RAM mobile data, AMPS, CT2, ERMES, land mobile, public safety
8PSK	Satellite, aircraft, telemetry pilots for monitoring broadband video systems
16 QAM	Microwave digital radio, modems, DVB-C, DVB-T
32 QAM	Terrestrial microwave, DVB-T
64 QAM	DVB-C, modems, broadband set top boxes, MMDS
256 QAM	Modems, DVB-C (Europe), Digital Video (US)

Amplitude-Shift Keying

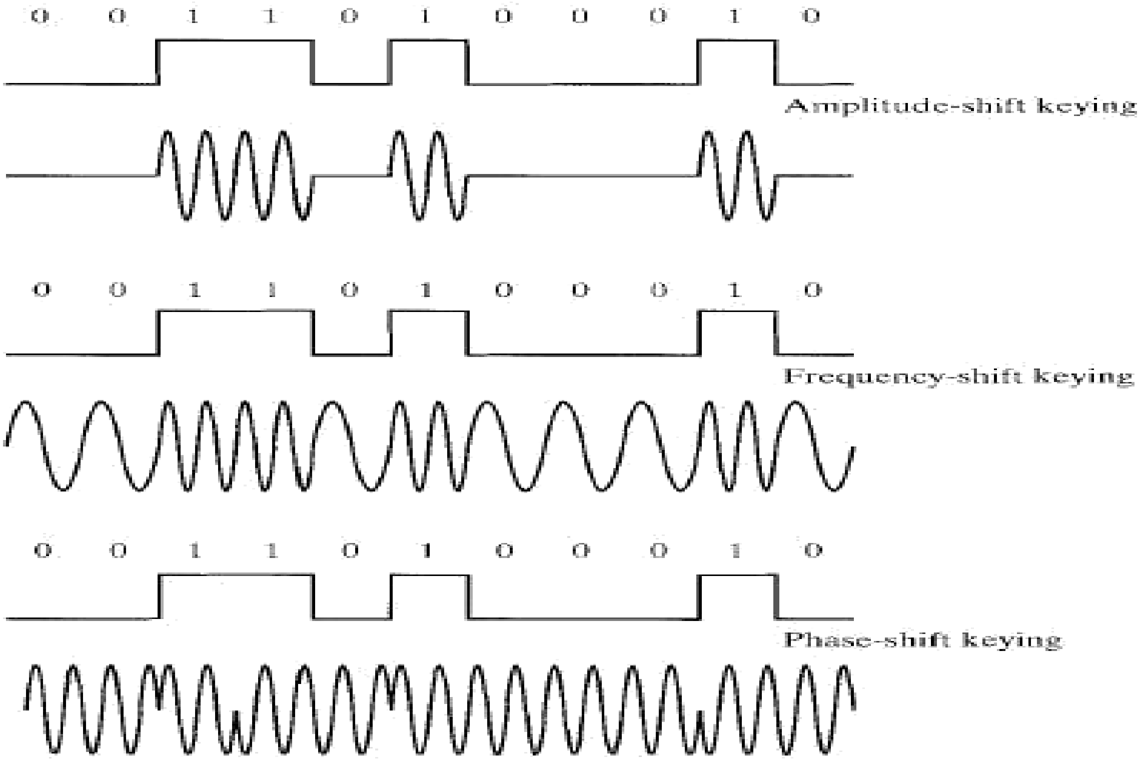
It is the simplest digital modulation technique where a binary information signal directly modulates the amplitude of an analog carrier. Only two output amplitudes are possible and ASK is sometimes called as digital amplitude modulation (DAM). Amplitude shift keying is given in mathematical terms as:

$$v_{ask}(t) = [\,1 + v_m(t)\,]\left[\,\frac{A}{2}\cos(\omega_c t)\right]$$

Where $v_{ask}(t)$ is amplitude-shift keying wave, $v_m(t)$ is digital modulation (modulating) signal in volts, $A/2$ is unmodulated carrier amplitude in volts and ω_c is analog carrier radian frequency in radians per second.



In the above equation, for the modulating signal $\mathbf{v_m(t)}$**,** logic 1 is represented by +1V and logic 0 is represented by -1V. So the modulated wave $v_{ask}(t)$ is either $\mathbf{A\cos(\omega_c t)}$ or 0 i.e., the carrier is either on or off. ASK is sometimes referred as on-off keying (OOK). The rate of change of the ASK waveform (baud) is the same as the rate of change of the binary input making bit rate equal to baud. With ASK, the bit rate is also equal to the minimum Nyquist bandwidth.



Modulation of analog signals for digital data.

Frequency Shift Keying

FSK is another simple, low-performance type of digital modulation. It is similar to FM, except the modulating signal is a binary signal varying between two discrete voltage levels. FSK is sometimes called as *binary* FSK (BFSK). FSK is generally expressed as

$$v_{fsk}(t) = \mathbf{V_c} \cos\{\, 2\pi[f_c + v_m(t)\Delta f]t \,\}$$

Where $v_{fsk}(t)$ is binary FSK waveform, $\mathbf{V_c}$ is peak analog carrier amplitude in volts, f_c is analog carrier center frequency in hertz, f is peak change or shift in the analog carrier frequency and $\mathbf{v_m(t)}$ is binary input(modulating) signal in volts. For logic 1, $\mathbf{v_m(t) = +1}$ and for logic 0, $\mathbf{v_m(t) = -1}$ reducing the equation to $v_{fsk}(t) = \mathbf{V_c} \cos\{\, 2\pi[f_c + f]t \,\}$ and $v_{fsk}(t) = \mathbf{V_c} \cos\{\, 2\pi[f_c - f]t \,\}$

As the binary signal changes from a logic 0 to a logic 1 and vice versa, the output frequency shifts between two frequencies: a mark, or logic 1 frequency (f_m) and a space or logic 0 frequency (f_s). The mark and space frequencies are separated from the carrier frequency by the peak frequency deviation (f) and from each other by 2 f.

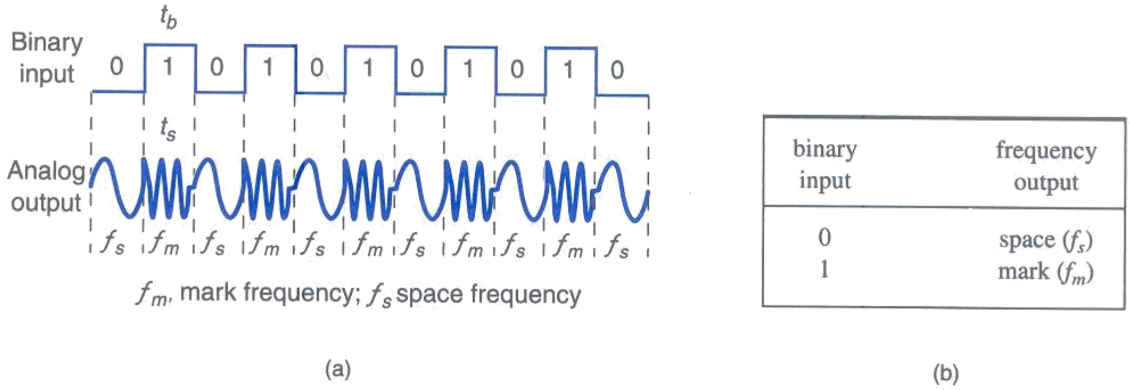


FIGURE 9-4 FSK in the time domain: (a) waveform; (b) truth table

With FSK, ***frequency deviation*** is defined as the difference between either the mark or space frequency and the center frequency or half the difference between the mark and space frequencies. Frequency deviation can be expressed as $f = |f_m - f_s| / 2$

The **baud** for BFSK is determined by placing N = 1, i.e., baud = $f_b/1 = f_b$

The **minimum bandwidth** for FSK is determined from;

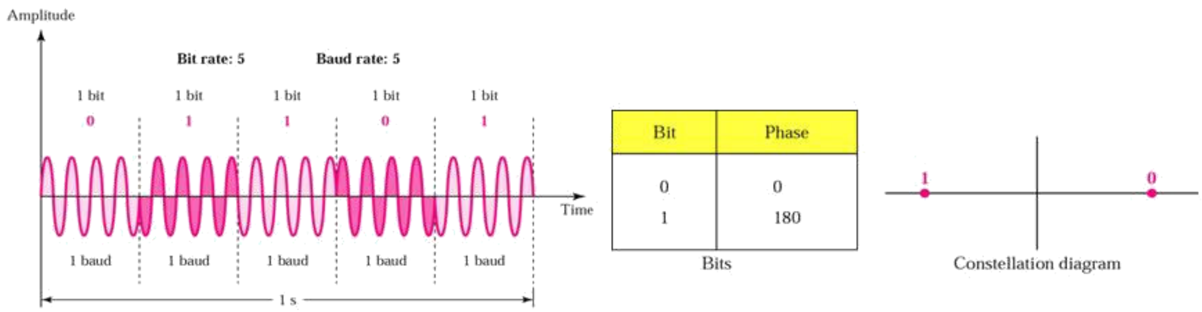
B = $|(f_s - f_b) - (f_m - f_b)| = |f_s - f_m| + 2f_b$. But $|f_s - f_m| = 2f$,

Therefore, **B = $2(f + f_b)$**, where B is minimum Nyquist bandwidth in hertz and f **is** frequency deviation and f_b is input bit rate.

Phase-Shift Keying

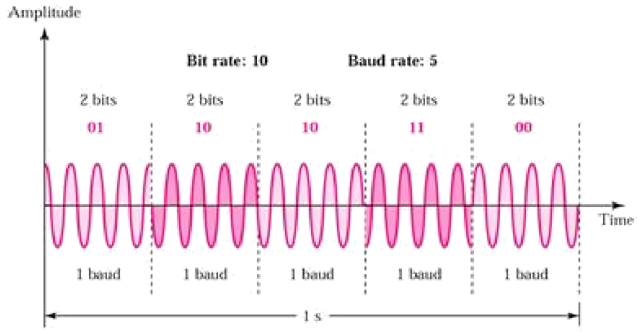
Phase-shift keying (PSK) is a digital modulation scheme that conveys data by changing, or modulating, the phase of a reference signal (the carrier wave). PSK uses a finite number of phases; each assigned a unique pattern of binary digits. Usually, each phase encodes an equal number of bits. PSK is not susceptible to the noise degradation that affects ASK or to the bandwidth limitations of FSK.

Binary phase-shift keying: The simplest PSK technique is called binary phase-shift keying (BPSK), where N = 1 and M = 2. Therefore, with BPSK two phases are possible for the carrier. It uses two opposite signal phases (0 and 180 degrees). The digital signal is broken up timewise into individual bits (binary digits). The state of each bit is determined according to the state of the preceding bit. If the phase of the wave does not change, then the signal state stays the same (0 or 1). If the phase of the wave changes by 180 degrees -- that is, if the phase reverses -- then the signal state changes (from 0 to 1 or from 1 to 0). Because there are two possible wave phases, BPSK is sometimes called ***biphase modulation or phase-reversal keying (PRK).***



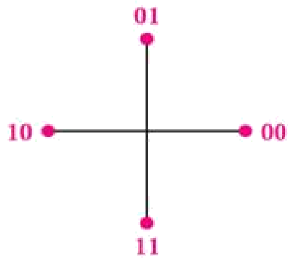
More sophisticated forms of PSK exist. In M-ary or multiple phase-shift keying (MPSK), there are more than two phases, usually four (0, +90, -90, and 180 degrees) or eight (0, +45, -45, +90, -90, +135, -135, and 180 degrees). If there are four phases (m = 4), the MPSK mode is called ***quadrature phase-shift keying*** or quaternary phase-shift keying (QPSK), and each phase shift represents two signal elements. If there are eight phases (m = 8), the MPSK mode is known as ***octal phase-shift keying (OPSK)***, and each phase shift represents three signal elements. In MPSK, data can be transmitted at a faster rate, relative to the number of phase changes per unit time, than is the case in BPSK.

QPSK is an M-ary encoding scheme where N = 2 and M = 4, which has four output phases are possible for a single carrier frequency needing four different input conditions. With two bits, there are four possible conditions: 00, 01, 10, and 11. With QPSK, the binary input data are combined into groups of two bits called ***dibits***.



Dibit	Phase
00	0
01	90
10	180
11	270

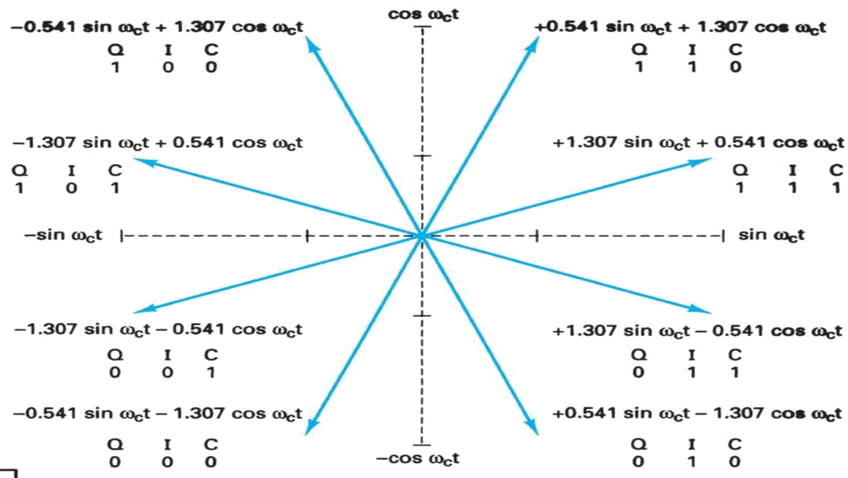
Dibit
(2 bits)



Constellation diagram

The above figure shows the output phase-versus-time relationship, truth table, and constellation diagram for QPSK. A phase of 0^0 now represents 00; 90^0 represents 01; 180^0 represents 10; and 270^0 represents 11. Data can be transmitted twice as efficiently using 4-PSK than 2-PSK.

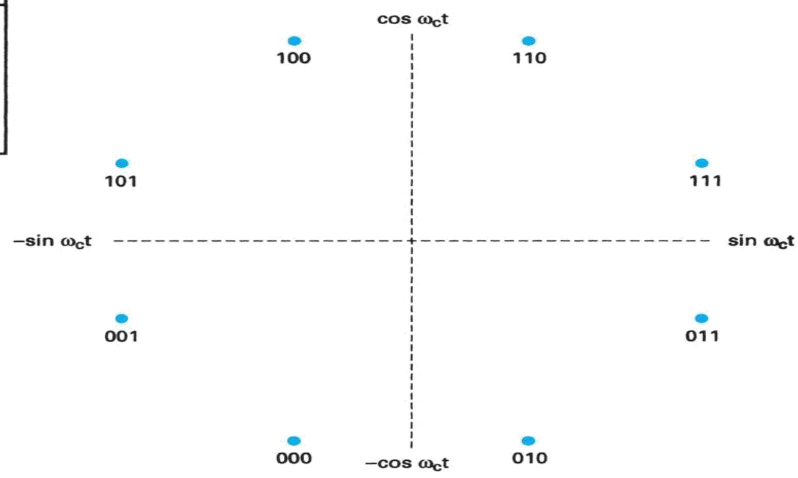
With 8-PSK, three bits are encoded forming ***tribits*** and producing eight different output phases. With 8-PSK, N = 3, M = 8, and the minimum bandwidth and baud equal one third the bit rate (f_b /3). 8-PSK is 3 times as efficient as 2-PSK.



(b)

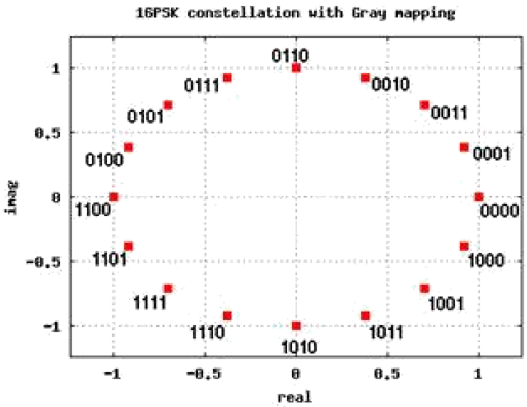
Binary input Q	I	C	8-PSK output phase
0	0	0	−112.5°
0	0	1	−157.5°
0	1	0	−67.5°
0	1	1	−22.5°
1	0	0	+112.5°
1	0	1	+157.5°
1	1	0	+67.5°
1	1	1	+22.5°

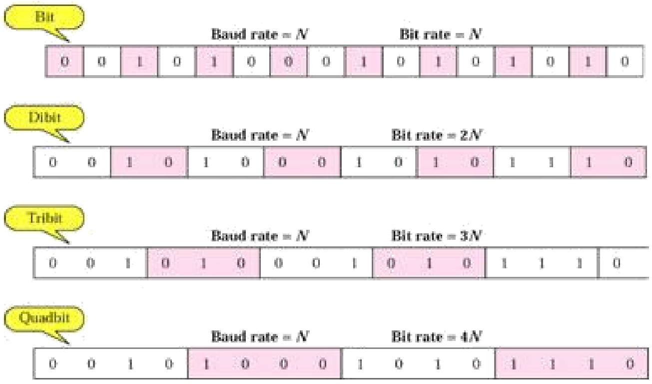
(a)



(c)

With 16-PSK, four bits called ***quadbits*** are combined, producing 16 different outputs phases. With 16-PSK, N = 4, M = 16, and the minimum bandwidth and baud equal one-fourth the bit rate (f_b /4).



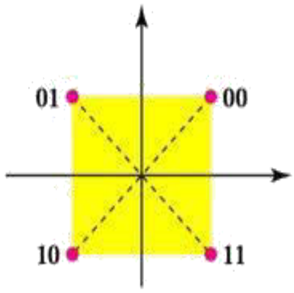


Modulation	Bit Rate	Encoding Scheme	Bandwidth Efficiency	Outputs Possible	Minimum Bandwidth	Baud
ASK	N	Single bit	1	2	f^{a}_{b}	f_b
FSK	N	Single bit	1	2	$>f_b$	f_b
BPSK	N	Single bit	1	2	f_b	f_b
QPSK	2N	Dibits	2	4	f_b /2	f_b /2
8-PSK	3N	Tribits	3	8	f_b /3	f_b /3
16-PSK	4N	Quadibits	4	16	f_b /4	f_b /4

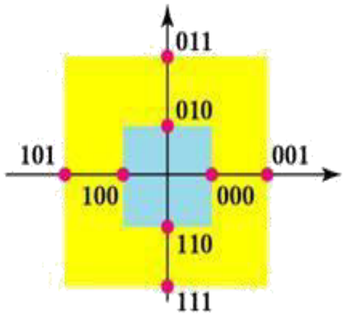
Quadrature Amplitude Modulation (QAM)

PSK is limited by the ability of the equipment to distinguish small differences in phase. Bandwidth limitations make combinations of FSK with other changes practically useless. Quadrature amplitude modulation is a combination of ASK and PSK so that a maximum contrast between each signal unit (bit, dibit, tribit, and so on) is achieved. QAM is used extensively as a modulation scheme for digital telecommunication systems. The primary advantage of QAM over PSK is immunity to transmission impairments, especially phase impairments that are inherent in all communication systems.

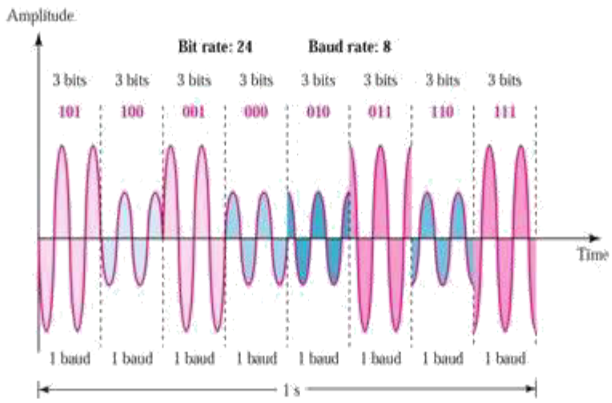
In 4-QAM and 8-QAM, number of amplitude shifts is fewer than the number of phase shifts. Because amplitude changes are susceptible to noise and require greater shift differences than do phase changes, the number of phase shifts used by a QAM system is always larger than the number of amplitude shifts.



4-QAM
1 amplitude, 4 phases



8-QAM
2 amplitudes, 4 phases



With 16-QAM, there are 12 phases and three amplitudes that are combined to produce 16 different output conditions. With QAM, there are always more phases possible than amplitude.

Bandwidth Efficiency

Bandwidth efficiency is often used to compare the performance of one digital modulation technique to another. It is the ration of transmission bit rate to the minimum bandwidth required for a particular modulation scheme. Mathematically represented as:

Bη = transmission bit rate (bps) / minimum bandwidth (Hz)

Modulation	Encoding Scheme	Outputs Possible	Minimum Bandwidth	Baud	Bη
ASK	Single bit	2	f_b	f_b	1
FSK	Single bit	2	f_b	f_b	1
BPSK	Single bit	2	f_b	f_b	1
QPSK	Dibits	4	$f_b/2$	$f_b/2$	2
8-PSK	Tribits	8	$f_b/3$	$f_b/3$	3
8-QAM	Tribits	8	$f_b/3$	$f_b/3$	3
16-PSK	Quadbits	16	$f_b/4$	$f_b/4$	4
16-QAM	Quadbits	16	$f_b/4$	$f_b/4$	4
32-PSK	Five bits	32	$f_b/5$	$f_b/5$	5
64-QAM	Six bits	64	$f_b/6$	$f_b/6$	6

Note: f_b indicates a magnitude equal to the input bit rate.

ASK, FSK, PSK, and QAM Summary

Trellis Code Modulation

Data Transmission Rates in excess of 56 kbps can be achieved over standard telephone circuits using an encoding scheme called trellis code modulation(TCM) developed by

Dr. Ungerboeck. It combines encoding and modulation to reduce the probability of error, thus improving the bit error performance and it uses conventional (tree) codes.

Trellis coding defines the manner in which signal-state transitions are allowed to occur, and transitions that do not follow this pattern are interpreted as transmission errors. TCM can improve error performance by restricting the manner in which signals are allowed to transition. TCM improves on standard QAM by increasing the distance between symbols on the constellation *(called Euclidean distance)*.

Appendix (some additional figures)

Binary input Q	Binary input I	QPSK output phase
0	0	−135°
0	1	−45°
1	0	+135°
1	1	+45°

(a)

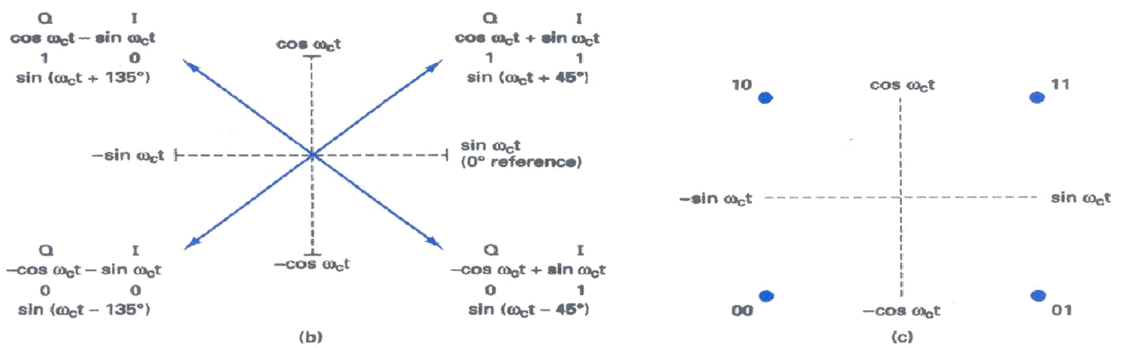


FIGURE 9-18 QPSK modulator: (a) truth table; (b) phasor diagram; (c) constellation diagram

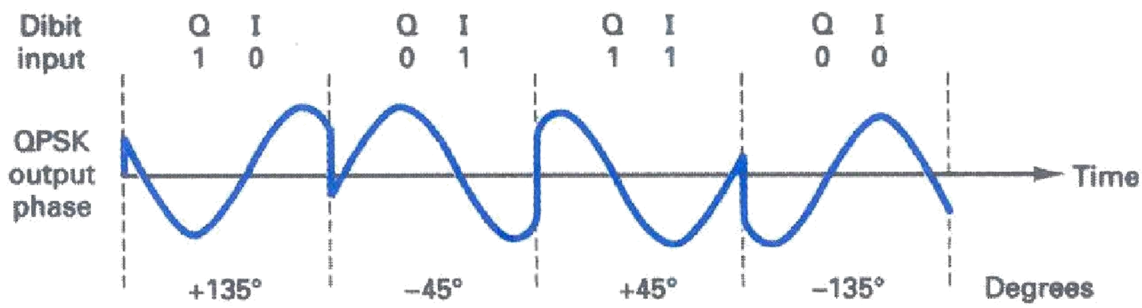
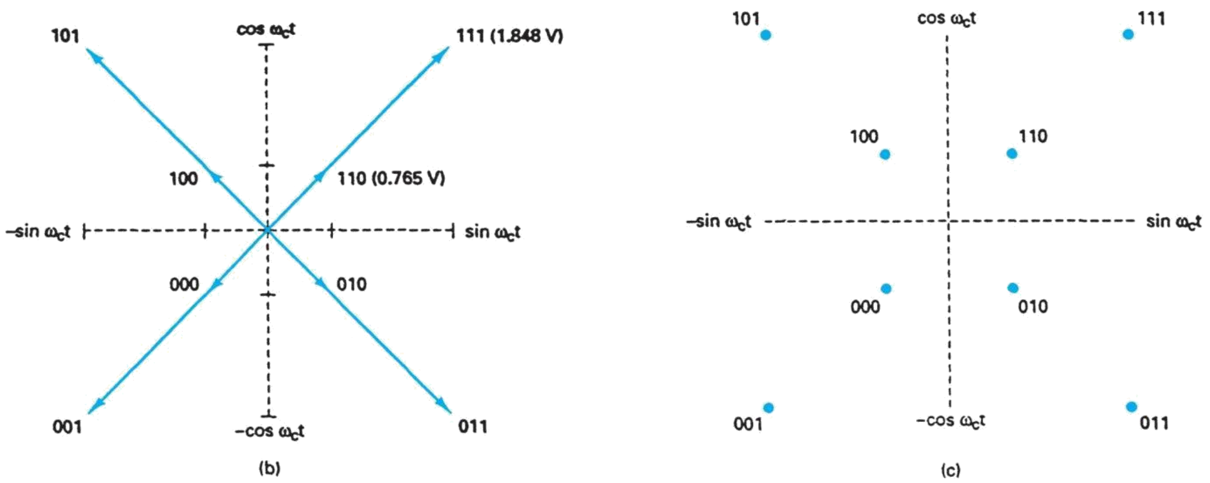


FIGURE 9-19 Output phase-versus-time relationship for a QPSK modulator

Binary input			8-QAM output	
Q	I	C	Amplitude	Phase
0	0	0	0.765 V	−135°
0	0	1	1.848 V	−135°
0	1	0	0.765 V	−45°
0	1	1	1.848 V	−45°
1	0	0	0.765 V	+135°
1	0	1	1.848 V	+135°
1	1	0	0.765 V	+45°
1	1	1	1.848 V	+45°

(a)



8-QAM modulator: (a) truth table; (b) phasor diagram; (c) constellation diagram

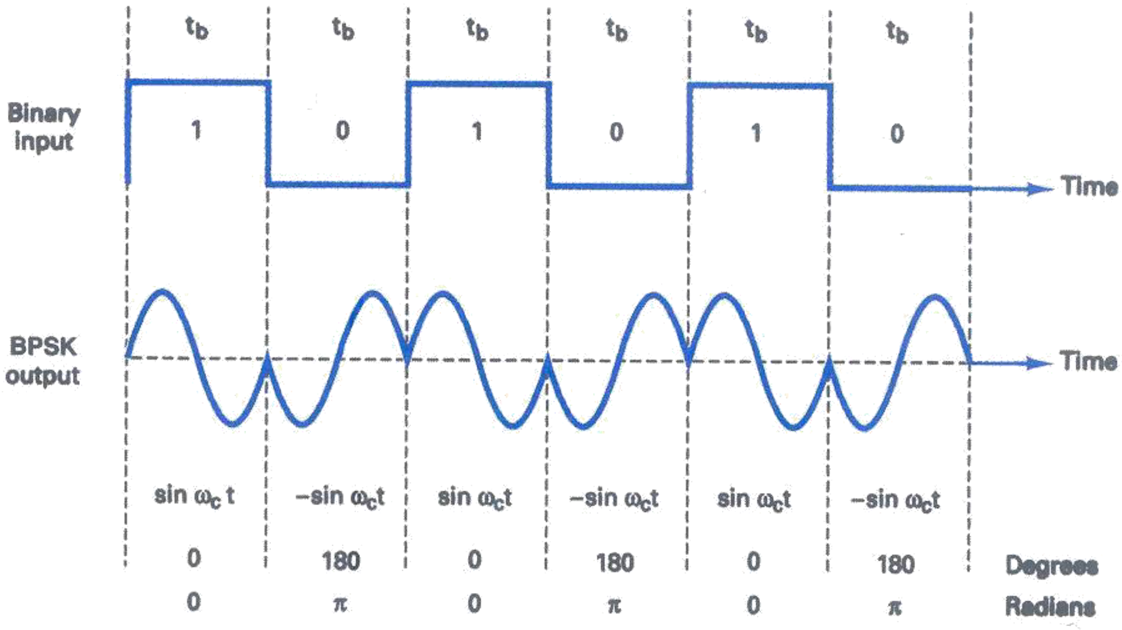


FIGURE 9-15 Output phase-versus-time relationship for a BPSK modulator

Assignment Questions

1. In QAM amplitude and phase of the transmitted signal are varied - Justify your answer with a block diagram and constellation diagram.

2. a. Explain the importance of asynchronous transmission in communication.
   b. Write the comparison between asynchronous and synchronous data transmission

3. (a) What is topology? Explain topologies in Data Communications?
   (b) What are the various types of transmission modes and explain.

4. (a) What is Data Communications? Explain briefly Data communication circuit.
   (b) Mention some standard organizations for Data Communications?

5. Draw OS I architectural model for open system inter networking and explain.

6. (a) Explain about Analog data, Digital signal encoding technique.
   (b) Differentiate between Data and Signals?

7. Sketch the binary ASK, FSK, PSK, and QPSK waveform for the following sequence 1011.

a) Explain the relationship between bits per second and baud for an FSK system.
b) Determine the bandwidth and baud for an FSK signal with a mark frequency of 24 kHz and a bit rate of 4 kbps.
c) Explain the relationship between
   i) Minimum bandwidth required for an FSK system and the bit rate
   ii) Mark and space frequencies

8. a) What is a constellation diagram? How it is used with PSK?
   b) Explain the minimum bandwidth required for a BPSK system and the bit rate.
   c) Explain M-ary.