Chapter 1

Introduction

Data Communication

Data communication is a vital part of the information society because it provides the infrastructure allowing the computers to communicate with one another. Data communications are the exchange of data between two devices via some form of transmission medium such as a wired or wireless. For data communications to occur, the communicating devices must be part of a communication system made up of combination of hardware and software. The effectiveness of a data communications system depends on four fundamental characteristics: **delivery, accuracy** and **timeliness.**

1. **Delivery:** The system must deliver data to the correct destination. Data must receive by the intended device or user and only by that device or user.
2. **Accuracy**: The system must deliver the data accurately. Data that had altered in transmission and left uncorrected are unusable.
3. **Timeliness**: The system must deliver data in a timely manner. Data delivered late are useless. Data delivering in the same order that they produced, and without significant delay. This kind of delivery is called real-time transmission.
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, let us assume that video packets is sent every 3D Ms. If some of the packets arrive with 3D-ms delay and others with 4D-ms de-lay, an uneven quality in the video is the result.

Data Representation

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

- **Text: -** In data communications, text is represented as a bit pattern, a sequence of bits (Os or 1s). Different sets of bit patterns have been design to represent text symbols. Each set a called a code, and the process of representing symbols is called coding. Today, the prevalent coding system is called Unicode, which uses 32 bits to represent a symbol or character used in any language in the world. The American Standard Code for Information Interchange (ASCII) developed some decades ago in the United States, now constitutes the first 127 characters in Unicode and referred to as Basic Latin.

- **Numbers: -** Numbers are also represented by bit patterns. However, a code such as ASCII is not use to represent numbers; the number directly converted to a binary number to simplify mathematical operations.

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

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

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

1.1 Evolution of Data Communications Systems

- **1837 - Samuel Morse** exhibited a working **Telegraph**, According to the History of computing has its earliest roots in telegraph systems. An account of data communication history posted by telecommunications experts at general telecom, LLC also points to a telegraph patent that inventor Charles Wheatstone filed that same year.
- **1843- Alexander Bain** patented a **Printing Telegraph**, Telegraph service used by the Great western Railway, an endorsement that allowed the service to expand across the nation.
- **1876 -** Improving on the telegraph, according to the History of Computing, **Alexander Graham Bell** introduce the **Telephone.**
- **1895 -** The development of early telecommunications **The Radio** is invention by **Guglielmo Marconi,** and it is make great developments in communication technology.
- **1947 - The Transistor** introduced by **Bells Labs,** This is a device that found integration in myriad subsequent electronic products.
- **1958** – These technologies are expanded by US government with its launch of a **Communications-oriented satellite,** and the first facsimile transmission over standard telephone lines occurred four years later.
- **1962 -** After the first **Fax Transmission,** the modulation of data into sound for transmission across telephone lines spread in popularity for several years. However, modulation/demodulation, or modem, technology continued to carry slower data traffic for the remainder of the 20th century.
- **1969 –** According to the History of Computing, the development of **Internet Protocol (IP)** marked a significant milestone in data communication history. Within the following decades, early packet communication technologies like ATM, Frame Relay and Integrated

Services Digital Network (ISDN) emerged as a viable solution for commercial and high-end residential data needs.

- **1991 –** More than 1 Million servers had come online using Internet Protocol Technology, and the World Wide Web emerged as the primary component of the Internet by the mid-1990s.

Basic Communication Model

Figure 1.1 Simple Communication model

The fundamental purpose of a communications system is the exchange of data between two parties. Figure presents one particular example, which is the communication between a workstation and a server over a public telephone network. Another example is the exchange of voice signals between two telephones over the same network. The key elements of the model are:

A Communication system has following components:

1. **Message:** It is the information or data to be share. It can consist of text, numbers, pictures, sound, video, or any combination of these.
2. **Sender**: The device/computer generates and sends that message.
3. **Receiver**: is the device or computer that receives the message. The location of receiver computer is generally different from the sender computer. The distance between sender and receiver depends upon the types of network used in between.
4. **Medium**: It is the channel or physical path through which the message carried from sender to the receiver. The medium can be wired as if twisted pair wire, coaxial cable, fiber-optic cable or wireless like laser, radio waves, and microwaves.
5. **Protocol**: It is a set of rules that is govern while the communication occur between the devices. Both sender and receiver follow same protocols to communicate with each other.

A protocol performs the following functions:

1. **Data sequencing**: It refers to breaking a long message into smaller packets of fixed size. Data sequencing rules define the method of numbering packets to detect loss or du-plication of packets, to identify correctly the packets, which belong to same message.
2. **Data routing**: Data routing defines the most efficient path between the source and destination.
3. **Data formatting**: Data formatting rules define which group of bits or characters within packet constitute data, control, addressing, or other information.
4. **Flow control**: A communication protocol also prevents a fast sender from overwhelming a slow receiver. It ensures resource sharing and protection against traffic congestion by regulating the flow of data on communication lines.

5. **Error control**: These rules designed to detect errors in messages and to ensure transmission of correct messages. The most common method is to retransmit erroneous message block. In such a case, a block having error discarded by the receiver and is re-transmitted by the sender.
6. **Precedence and order of transmission**: These rules ensure that all the nodes get a chance to use the communication lines and other resources of the network based on the priorities assigned to them.
7. **Connection establishment and termination**: These rules define how connections are established, maintained and terminated when two nodes of a network want to communicate with each other.
8. **Data security**: Providing data security and privacy built into most communication software packages. It prevents access of data by unauthorized users.
9. **Log information**: Several communication software designed to develop log information, which consists of all jobs and data communications tasks that have taken place. Such information may be use for charging the users of the network based on their usage of the network resources.

1.2 Analog and Digital Data Transmission, Data Communication Terminology

An **Analog signal** is any continuous signal for which the time varying feature (variable) of the signal is a representation of some other time varying quantity, i.e., analogous to another time varying signal. It differs from a digital signal in terms of small fluctuations in the signal, which are meaningful.

A **digital signal** uses discrete (discontinuous) values. By contrast, non-digital (or analog) systems use a continuous range of values to represent information. Although digital representations are discrete, the information represented can be either discrete, such as numbers or letters, or continuous, such as sounds, images, and other measurements of continuous systems.

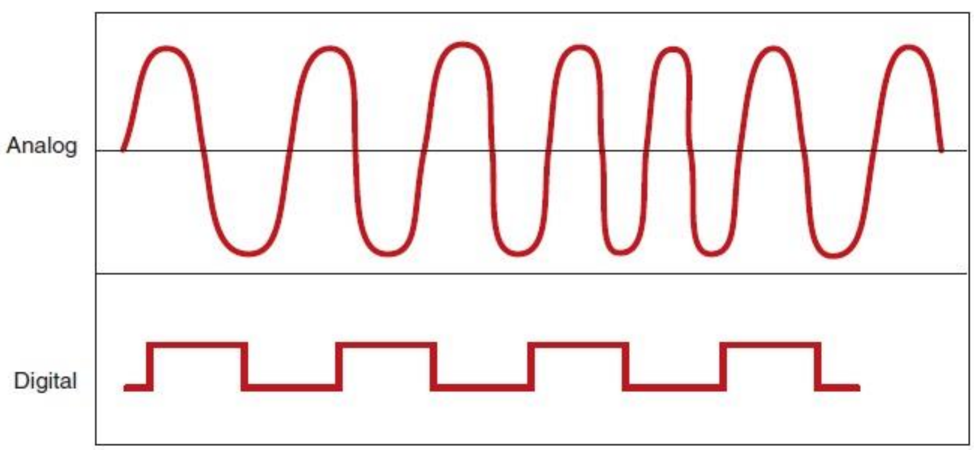


Figure1.2: Analog and Digital Signal

Analog and digital signals used to transmit information, usually through electric signals. In both these technologies, the information, such as any audio or video, transformed into electric signals. The difference between analog and digital technologies is that in analog technology, information translated into electric pulses of varying amplitude. In digital technology, translation of information is into binary format (zero or one) where each bit is representative of two distinct amplitudes.

Properties of Digital vs Analog signals

Digital information has certain properties that distinguish it from analog communication methods. These include

- Synchronization – digital communication uses specific synchronization sequences for determining synchronization.
- Language – digital communications requires a language, which should be possess by both sender and receiver and should specify meaning of symbol sequences.
- Errors – disturbances in analog communication causes errors in actual intended communication but disturbances in digital communication does not cause errors enabling error free communication. Errors should be able to substitute, insert or delete symbols to be expressed.
- Copying – analog communication copies are quality wise not as good as their originals while due to error free digital communication, copies can be made indefinitely.
- Granularity – for a continuously variable analog value to be represented in digital form there occur quantization error which is difference in actual analog value and digital representation and this property of digital communication is known as granularity.

Differences in Usage in Equipment

Many devices come with built in translation facilities from analog to digital. Microphones and speaker are perfect examples of analog devices. Analog technology is cheaper but there is a limitation of size of data that can be transmitted at a given time.

Digital technology has revolutionized the way most of the equipment's work. Data is converted into binary code and then reassembled back into original form at reception point. Since these can be easily manipulated, it offers a wider range of options. Digital equipment is more expensive than analog equipment.

Comparison of Analog vs Digital Quality

Digital devices translate and reassemble data and in the process are more prone to loss of quality as compared to analog devices. Computer advancement has enabled use of error detection and error correction techniques to remove disturbances artificially from digital signals and improve quality.

Differences in Applications

Digital technology has been most efficient in cellular phone industry. Analog phones have become redundant even though sound clarity and quality was good.

Analog technology comprises of natural signals like human speech. With digital technology, this human speech can be saved and stored in a computer. Thus, digital technology opens up the horizon for endless possible uses.

<u>Data Communication Terminology</u>

- Data Channel: -In this medium data carries from one point to another point.
- Baud: - Each communication channel has certain capacity and it can carry information up to that extent only. This capacity measured in terms of Baud.
- Bits per Second: -The speed at which data transferred between two points measured in terms of Bits per Second or bps.
    - bps – bits per second,
    - Bps – Bytes per second (Note capital B)
- Bandwidth – The amount of data a communication system can transfer per unit time referred as Bandwidth of the system. Bandwidth simply indicates the data transfer rate. The more the data needed to be transmitted in the given unit time the more should be the bandwidth. Alternatively, it can be said that more the bandwidth of the communication system more will be the data transfer rate. Bandwidth measured in bps or Baud. Generally, a Baud is identical to bits per second. A rate of 100 Baud is equal to 100 bps.
- In digital context, the level of bandwidth falls into three category:
    - Narrowband – Speed of narrowband varies between 45 to 300 Baud. Low speed devices use this narrowband channels.
    - Voice-band- Speed of voice-band channels ranges up to 9600 Baud. They are generally use in the ordinary telephone voice communication.
    - Broadband – The speed of broadband channels ranges up to 1 million Baud or more. High-speed devices use broadband for large volume of data transfer at high rate. Broadcast television, microwave and satellite uses broadband channel.
- Data Transfer Rates: - The amount of data transferred per second by a communication channel is known as data transfer rate. It is measure in bits per second (bps).

1.3 Standard Organizations

An association of organizations, governments, manufacturers and users form the standards organizations and are responsible for developing, coordinating and maintaining the standards .The purpose is that all data communications equipment are 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 comprised 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 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) 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 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.

Applications

- Airline reservation system
  - American airline: Sabre system
  - United airline: Apollo reservation system
- Automatic teller machine
  - Swift: Society for Worldwide Interbank Financial Telecommunication
- Sales order entry
  - Point of sale
  - Universal product code
- Unstructured data application
  - Electronic mail
    - Ownership of content
  - Simple mail transfer protocol (SMTP)
    - No foreign characters
    - No executable files
    - Limited size
  - Multipurpose Internet mail extensions (MIME)

****

Chapter 2

Data Transmission

Data Transmission

Data transmission refers to the process of transferring data between two or more digital devices. Data is transmitted from one device to another in analog or digital format. Basically, data transmission enables devices or components within devices to speak to each other.

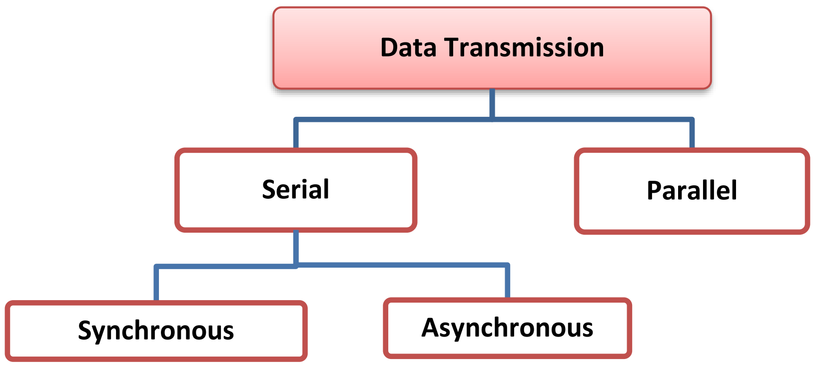


Figure 2.1: Data Transmission Classification

2.1 Parallel and Serial Transmission

How does data transmission work between digital devices?

Data is transferred in the form of bits between two or more digital devices. There are two methods used to transmit data between digital devices: serial transmission and parallel transmission. Serial data transmission sends data bits one after another over a single channel. Parallel data transmission sends multiple data bits at the same time over multiple channels.

Serial Transmission

When data is sent or received using serial data transmission, the data bits are organized in a specific order, since they can only be sent one after another. The order of the data bits is important as it dictates how the transmission is organized when it is received. It is viewed as a reliable data transmission method because a data bit is only sent if the previous data bit has already been received.

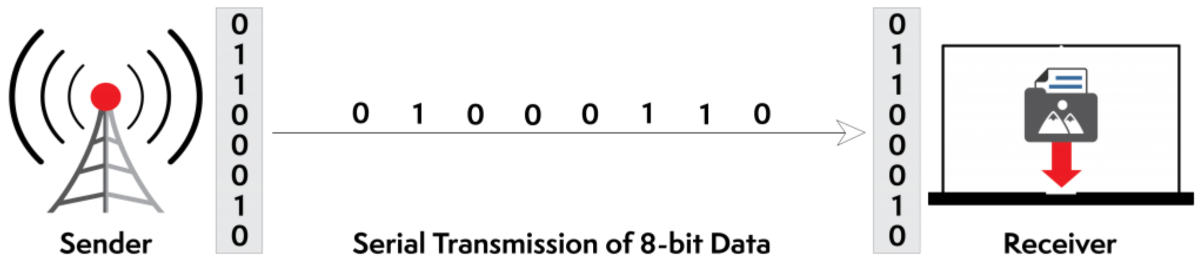


Figure 2.2: Serial Data Transmission

Serial transmission has two classifications: **asynchronous and synchronous.**

1. <u>Asynchronous Serial Transmission</u>

Data bits can be sent at any point in time. Stop bits and start bits are used between data bytes to synchronize the transmitter and receiver and to ensure that the data is transmitted correctly. The time between sending and receiving data bits is not constant, so gaps are used to provide time between transmissions.

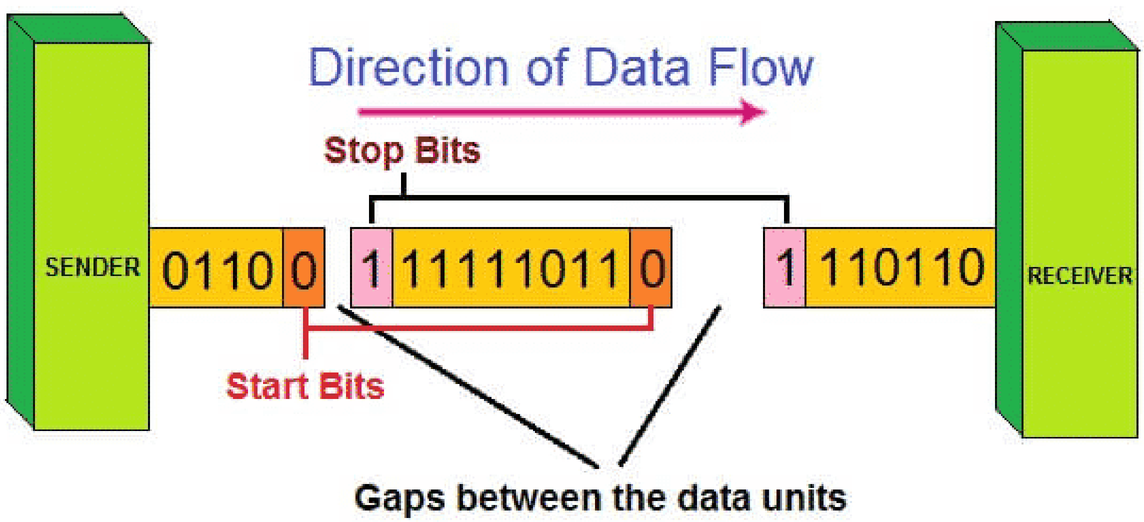


Figure2.3: Asynchronous Serial Transmission

The advantage of using the asynchronous method is that no synchronization is required between the transmitter and receiver devices. It is also a more cost effective method. A disadvantage is that data transmission can be slower, but this is not always the case.

Advantages of Asynchronous transmission

i. This method of data transmission is cheaper in cost as compared to synchronous e.g. If lines are short, asynchronous transmission is better, because line cost would be low and idle time will not be expensive.

ii. In this approach each individual character is complete in itself, therefore if character is corrupted during transmission, its successor and predecessor character will not be affected.
iii. It is possible to transmit signals from sources having different bit rates.
iv. The transmission can start as soon as data byte to be transmitted becomes available.
v. Moreover, this mode of data transmission in easy to implement.

Disadvantages of asynchronous transmission

i. This method is less efficient and slower than synchronous transmission due to the overhead of extra bits and insertion of gaps into bit stream.
ii. Successful transmission inevitably depends on the recognition of the start bits. These bits can be missed or corrupted.

2. <u>Synchronous Serial Transmission</u>

Data bits are transmitted as a continuous stream in time with a master clock. The data transmitter and receiver both operate using a synchronized clock frequency; therefore, start bits, stop bits, and gaps are not used. This means that data moves faster and timing errors are less frequent because the transmitter and receiver time is synced. However, data accuracy is highly dependent on timing being synced correctly between devices. In comparison with asynchronous serial transmission, this method is usually more expensive.

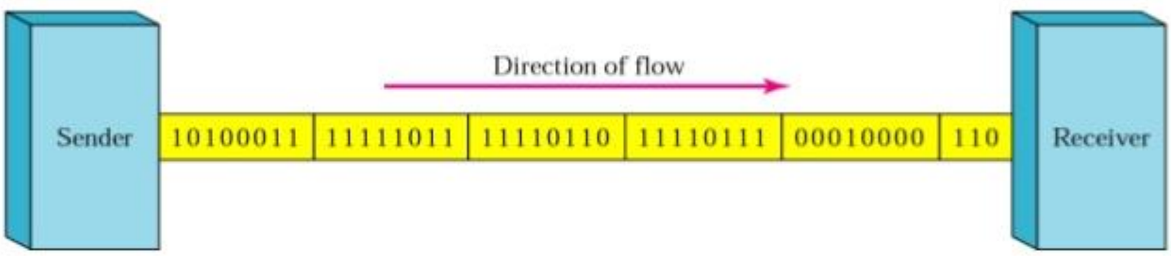


Figure 2.4: Synchronous serial transmission

Application of Synchronous transmission

i. Synchronous transmission is used for high speed communication between computers.

Advantage of Synchronous transmission

i. This method is faster as compared to asynchronous as there are no extra bits (start bit & stop bit) and also there is no gap between the individual data bytes.

Disadvantages of Synchronous transmission

i. It is costly as compared to asynchronous method. It requires local buffer storage at the two ends of line to assemble blocks and it also requires accurately synchronized clocks at both ends. This lead to increase in the cost.
ii. The sender and receiver have to operate at the same clock frequency. This requires proper synchronization which makes the system complicated.

When is serial transmission used to send data?

Serial transmission is normally used for long-distance data transfer. It is also used in cases where the amount of data being sent is relatively small. It ensures that data integrity is maintained as it transmits the data bits in a specific order, one after another. In this way, data bits are received in-sync with one another.

Parallel Transmission

When data is sent using parallel data transmission, multiple data bits are transmitted over multiple channels at the same time. This means that data can be sent much faster than using serial transmission methods.

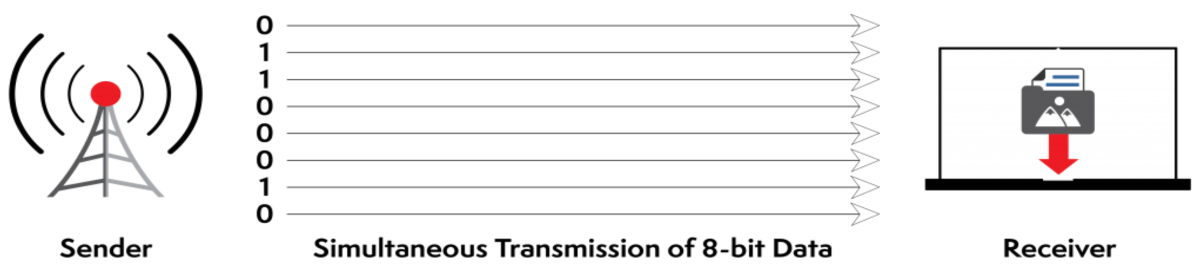


Figure 2.5: Parallel Transmission

Given that multiple bits are sent over multiple channels at the same time, the order in which a bit string is received can depend on various conditions, such as proximity to the data source, user location, and bandwidth availability. Two examples of parallel interfaces can be seen below. In the first parallel interface, the data is sent and received in the correct order. In the second parallel interface, the data is sent in the correct order, but some bits were received faster than others.

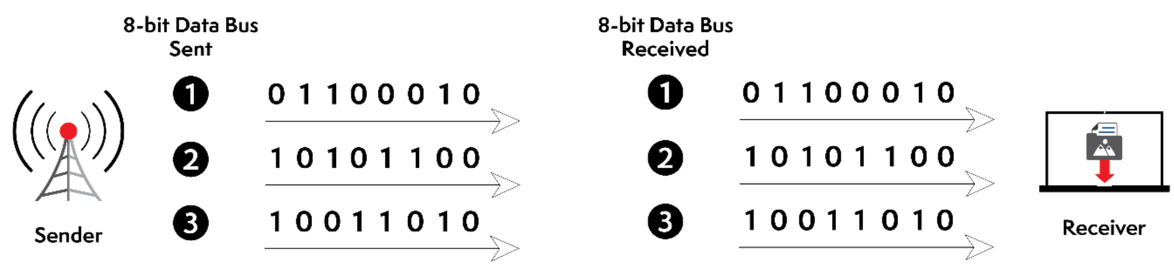


Figure 2.6: Example of Parallel Transmission – Data Received Correctly

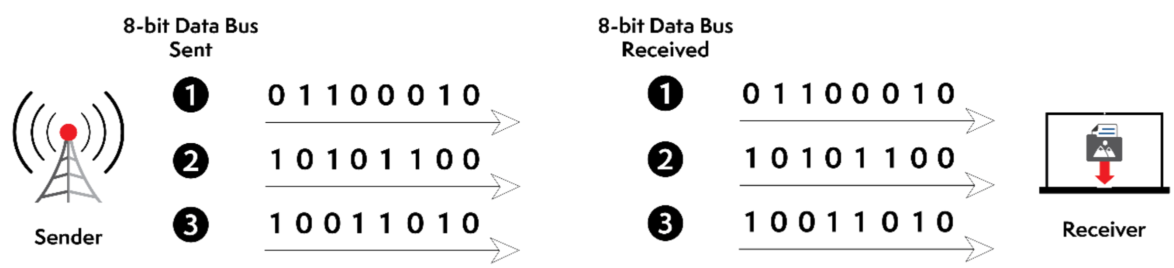


Figure 2.7: Example of Parallel Transmission – Data Received Incorrectly

Advantages and Disadvantages of Using Parallel Data Transmission

The main advantages of parallel transmission over serial transmission are:

- It is easier to program.
- Data is sent faster.

Although parallel transmission can transfer data faster, it requires more transmission channels than serial transmission. This means that data bits can be out of sync, depending on transfer distance and how fast each bit loads. A simple of example of where this can be seen is with a voice over IP (VOIP) call when distortion or interference is noticeable. It can also be seen when there is skipping or interference on a video stream.

When is parallel transmission used to send data?

Parallel transmission is used when:

- A large amount of data is being sent.
- The data being sent is time-sensitive.
- And the data needs to be sent quickly.

A scenario where parallel transmission is used to send data is video streaming. When a video is streamed to a viewer, bits need to be received quickly to prevent a video pausing or buffering. Video streaming also requires the transmission of large volumes of data. The data being sent is also time-sensitive as slow data streams result in poor viewer experience.

2.2 Line Configuration

A link is a communication pathway that transfer data from one device to another. Devices can be a computer, printer or any other device that is capable to send and receive data. For visualization purpose, imagine any link as a line drawn between two points.

For communication to occur, two devices must be connected in some way to the same link at the same time. There are two possible types of connections:

- Point-to-Point Connection
- Multipoint Connection

Point-to-Point Connection :

A point-to-point connection provides a dedicated link between two devices. The entire capacity of the link is reserved for transmission between those two devices. Most point-to-point connections use a actual length of wire or cable to connect the two end, but other options such as microwave or satellite links are also possible. Point to point network topology is considered to be one of the easiest and most conventional network topologies. It is also the simplest to establish and understand. e.g TV and Remote.

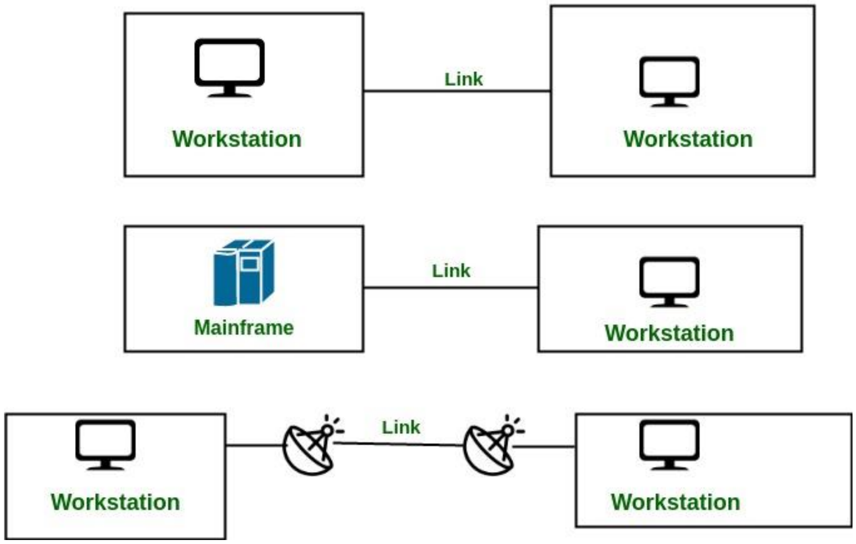


Figure 2.8: Point to Point Protocol

Multipoint Connection :

It is also called Multidrop configuration. In this connection two or more devices share a single link. More than two devices share the link that is the capacity of the channel is shared now. With shared capacity, there can be two possibilities in a Multipoint Line configuration:

- **Spatial Sharing:** If several devices can share the link simultaneously, its called Spatially shared line configuration.
- **Temporal (Time) Sharing:** If users must take turns using the link , then its called Temporally shared or Time Shared Line configuration.

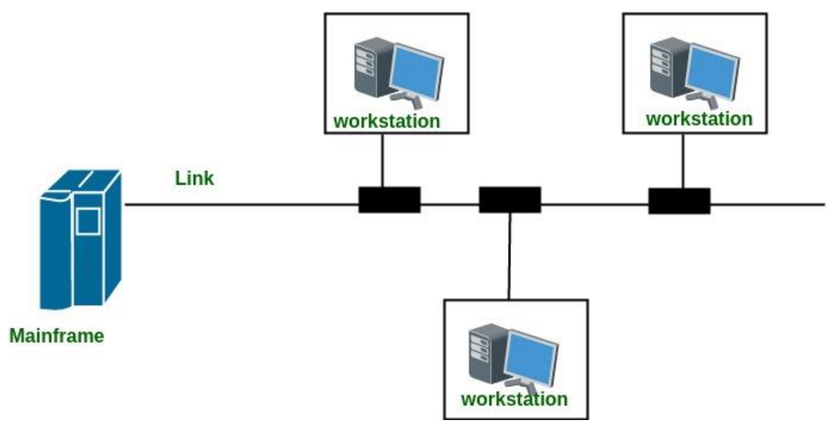


Figure 2.9: Multipoint Protocol (Temporal Sharing)

2.3 Bit Rate/ Baud Rate, Transmission Channel, RS-232C and RS-449 Interface Standards

Bit Rate/Baud Rate

Bit rate and Baud rate are generally used in data communication,

Bit rate is the transmission of number of bits per second. On the other hand, Baud rate is defined as the number of signal units per second. The formula which relates both bit rate and baud rate is given below:

Bit rate = Baud rate x the number of bit per baud.

Difference between bit rate and baud rate

1. Bit rate is defined as the transmission of number of bits per second where as Baud rate is defined as the number of signal units per second.
2. Bit rate is also defined as per second travel number of bits where as Baud rate is also defined as per second number of changes in signal.
3. Bit rate emphasized on computer efficiency While baud rate emphasized on data transmission.
4. The formula of Bit Rate is:- baud rate x the number of bit per baud
5. The formula of Baud Rate is:- bit rate / the number of bit per baud

Transmission Channels

The data transmission channels is needed for data transmission. These can be **Guided Channels** or **Unguided Channels.**

Guided Media:

Guided media use a physical connection between two devices. The waves are guided along a physical path over the medium. A signal has to travel within the physical limits of the guided medium. These may be:

- **Twisted Pair Cable**: Twisted pair cable is a cable, which is made by intertwining two separate insulated wires together. A twisted pair cable consists of two conductors, which are normally made of copper. This pair has a bandwidth to distance ratio of about 1 MHz per kilometer. These are of two types: shielded and un-shielded. Shielded Twisted Pair (STP) has a fine wire mesh surrounding the wires, which helps to protect the transmission Unshielded Twisted Pair (UTP) does not have, that mesh. Shielded cable used in older telephone networks, network, and data communications to reduce outside interference.

- **Optical Fiber:** Optical fiber cable is a cable made of optical fibers that can transmit large amounts of information at the speed of light. It is a medium and the technology associated with the transmission of information as light pulses along a glass or plastic strand. Optical fiber carries much more information than conventional copper wire. In one cable, there are many optical fibers. The glass fiber requires more protection within an outer cable than copper. It consists of two concentric cylinders: an inner core and a cladding used to surround the core. These are made of transparent plastic or glass material. The density of the core and cladding must be such to be helpful to reflect the beam of light. The core guides the beam and cladding prevents it. A laser or Light Emitting Diode (LED) usually generates the signal. The speed of data transfer is high. Due to high speed and little disturbance, these being used rapidly in telecommunications.

Advantages of Optical Fiber

- Higher Bandwidth helps data at a higher rate.
- Less signal attenuation.
- Immunity to electromagnetic interface.
- Lightweight.
- Limitations of Optical Fiber
- It needs experts to install it and provide maintenance.
- Being unidirectional needs two fibers for bi-directional communication.
- Cost factor is high.

Unguided Media

Unguided media transfers the signals through wireless medium. It transports electromagnetic waves in the air, which received by the devices to catch them. These can be:

- **Radio Waves:** Radio waves are an invisible form of electromagnetic radiation, and it is one of the widest ranges in the electromagnetic spectrum. A radio wave is an electromagnetic wave propagated by an antenna. The frequency ranges from 3 Kilo-Hertz to 1 Giga-Hertz. These can travel in any direction and easy to produce waves. It eliminates the cost of physical medium and is very useful for long distanced communication. These are of a very long wavelength, such as thousands of meters, tend to travel along the surface of the earth and even penetrate into the water. These are useful for communication with submarines and for broadcasting time signals, radio broadcast, cellular telephones, etc. Each communication service uses a part of the spectrum that is suitable for its needs. These days' cellular radios are used to provide mobile phone networks. These operate in the VHF (Very High Frequency) band. These help in multi casting, which means it, transmit a signal for specific group, which may be more than one. These used in FM radios, cordless phones, etc. A communication between single source and destination known as unicast and if there are many receivers to catch the signals sent by sender at any destination, it is called broadcast.

- **Bluetooth**: Bluetooth is used to send and receive data over short distance in mobile and related technology. A Bluetooth connection is wireless and automatic, and it has a number of interesting features that can simplify our daily lives.

- **Microwaves**: Microwaves are radio frequencies, which ranges between about 1 GHz (one gigahertz) to about 300 GHz. It may be defined as a short electromagnetic wave (longer than infrared but shorter than radio waves) used for radar and ovens and for transmitting telephone, facsimile video and data. These are unidirectional waves and hence have less interference by a pair of aligned antenna to another. Uses of Microwave are radio transmission, telecommunication carriers and TV stations.

  Satellite is another form of microwave system. Repeaters present in the sky supplement it. Satellites have a high bandwidth and can support variety of channels. It has some limitations such as:

    - High set up cost.
    - The lifetime is limited.

  - These waves cannot be receive inside the building.

RS-232C

- RS-232C is a long-established standard ("C" is the current version) that describes the physical interface and protocol for relatively low-speed serial data communication between computers and related devices.
- An industry trade group, the Electronic Industries Association (EIA), defined it originally for teletypewriter devices.
- RS-232C is the interface that your computer uses to talk to and exchange data with your modem and other serial devices.
- Somewhere in your PC, typically on a Universal Asynchronous - Receiver/Transmitter (UART) chip on your motherboard, the data from your computer is transmitted to an internal or external modem (or other serial device) from its Data Terminal Equipment (DTE) interface.
- Since data in your computer flows along parallel circuits and serial devices can handle only one bit at a time, the UART chip converts the groups of bits in parallel to a serial stream of bits.
- As your PC's DTE agent, it also communicates with the modem or other serial de-vice, which, in accordance with the RS-232C standard, has a complementary inter-face called the Data Communications Equipment (DCE) interface.
- One of the advantages of a serial system is that it lends itself to transmission over tele-phone lines.
- The serial digital data can be converted by modem, placed onto a standard voice-grade telephone line, and converted back to serial digital data at the receiving end of the line by another modem.
- Officially, RS232 is defined as the "Interface between data terminal equipment and data communication equipment using serial binary data exchange."
- This definition defines data terminal equipment, as the equipment is the modem.
- A modem cable has pin-to-pin connections and designed to connect DTE device to a DCE device.
- The RS-232C standard is an asynchronous serial communication method.
- Serial means that the information is sent 1-bit at a time.
- Asynchronous means that no clock signal is sent with the data. Each side uses its own clock and a start and stop bit. Synchronous communication means that a clock signal is sent in addition to a data signal.
- The RS-232C standard works at the physical layer of the communication standard. This is the lowest level and the one that physically connects the devices.
- The communication is done through the serial port of the PC. This is a male connector with 25 (old) or 9 (new) pins, in both cases only 9 pins, at the most, are used.

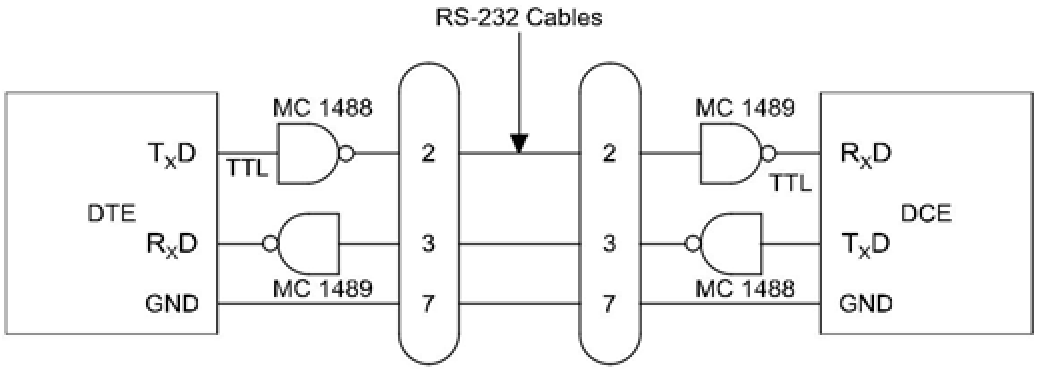


Figure 2.30: Connection between DCE to DTE

Interface Standards:

- In telecommunications, an interface standard is a standard that describes one or more functional characteristics (such as code conversion, line assignments, or protocol compliance) or physical characteristics (such as electrical, mechanical, or optical characteristics) necessary to allow the exchange of information between two or more (usually different) systems or pieces of equipment.
- An interface standard may include operational characteristics and acceptable levels of performance.
- In the military community, interface standards permit command and control functions to be performed using communication and computer systems.

RS 449 Interface

- Is able to send data at high Speed without stray noise causing interference is to use a differential form of signaling.
- As the RS449 receivers use a differential input, and they are not reference to ground, any noise that is pick up does not affect the input.
- This means that higher levels of noise can be tolerate without any degradation to the performance to the data communications system.
- For the RS449 interface, ten additional circuits' functions have been provided when compared to RS232.
- In addition to this the RS449 interface requires the use of 37 way D-type connectors and 9 way D-type connectors, the latter being necessary when use is made of the secondary channel interchange circuits.

Transmission Modes

Transmission mode means transferring of data between two devices. It is also known as communication mode. Buses and networks are design to allow communication to occur between individual devices that are interconnected. There are three types of transmission mode:

- Simplex Mode
- Half-Duplex Mode
- Full-Duplex Mode

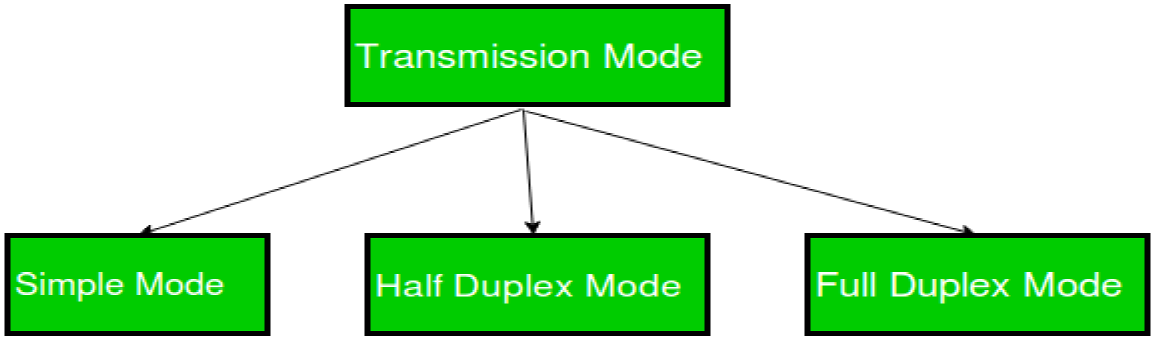


Figure 4.11: Types of Transmission Modes

- ***Simplex Mode***

  In Simplex mode, the communication is unidirectional, as on a one-way street. Only one of the two devices on a link can transmit, the other can only receive. The simplex mode can use the entire capacity of the channel to send data in one direction.

  Example: Keyboard and traditional monitors. The keyboard can only introduce input; the monitor can only give the output.

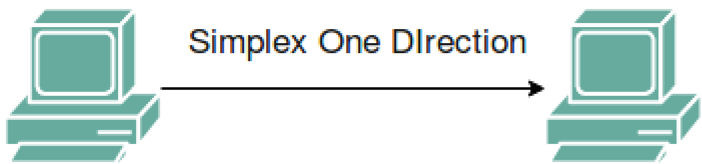


Figure 2.12: Simplex

- ***Half-Duplex Mode***

  In half-duplex mode, each station can both transmit and receive, but not at the same time. When one device is sending, the other can only receive, and vice versa. The half-duplex mode is used in cases where there is no need for communication in both direction at the same time. The entire capacity of the channel can be utilized for each direction. Example: Walkie- talkie in which message is sent one at a time and messages are sent in both the directions.

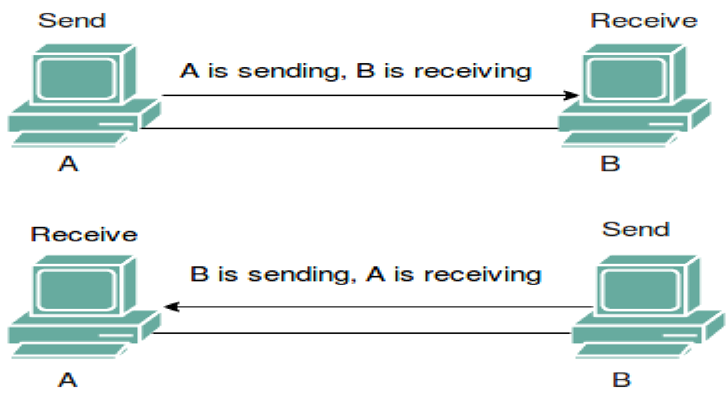


Figure 2.15: Half- Duplex

- ***Full-Duplex Mode***

  In full-duplex mode, 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 other direction, this sharing can occur in two ways:

  - Either the link must contain two physically separate transmission paths, one for sending and other for receiving.
  - Alternatively, the capacity is divided between signals travelling in both directions.

Full-duplex mode is use when communication in both direction is required all the time. The capacity of the channel however must be divided between the two directions.

Example: Telephone Network in which there is communication between two persons by a telephone line, through which both can talk and listen at the same time.

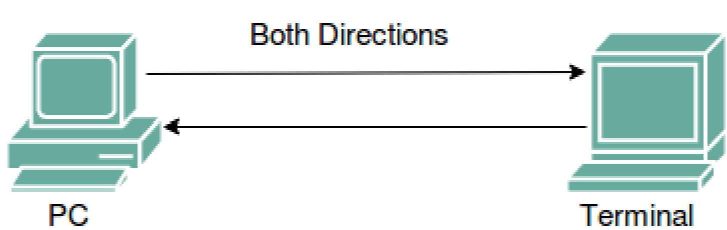


Figure2.16: Full Duplex

****

Chapter 3

Signals and Systems

Signals

The signal. Stated in mathematical terms, a signal is merely a function. Analog signals are continuous-valued; digital signals are discrete-valued. The independent variable of the signal could be time (speech, for example), space (images), or the integers (denoting the sequencing of letters and numbers in the football score).

Signals are classified into the following categories:

- Continuous Time and Discrete Time Signals
- Deterministic and Non-deterministic Signals
- Even and Odd Signals
- Periodic and Aperiodic Signals
- Energy and Power Signals
- Real and Imaginary Signals

i. Continuous Time and Discrete Time Signals

A signal said to be continuous when it is define for all instants of time.

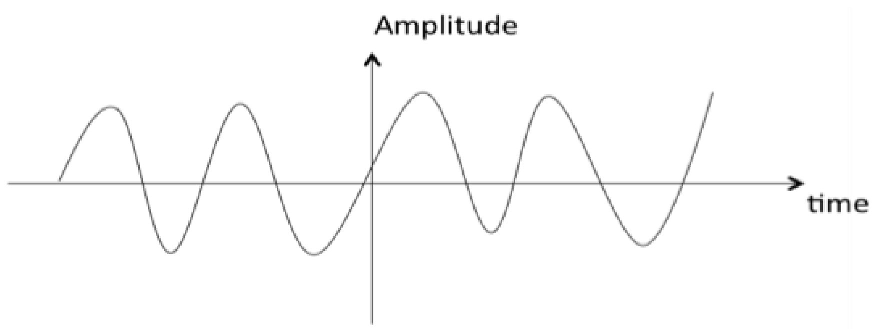


Figure3.7: Continuous Signal

A signal is said to be discrete when it is defined at only discrete instants of time/

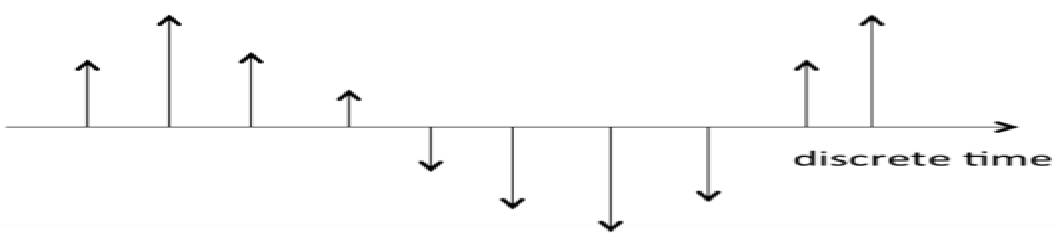


Figure3.8: Discrete Signal

ii. Deterministic and Non-deterministic Signals

A signal is said to be deterministic if there is no uncertainty with respect to its value at any instant of time. Or, signals which can be defined exactly by a mathematical formula are known as deterministic signals.

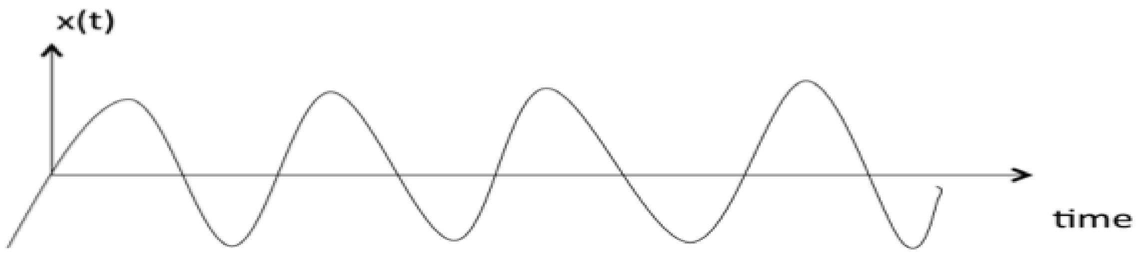


Figure 3.9: Deterministic Signal

A signal is said to be non-deterministic if there is uncertainty with respect to its value at some instant of time. Non-deterministic signals are random in nature hence they are called random signals. Random signals cannot be describe by a mathematical equation. They are model in probabilistic terms.

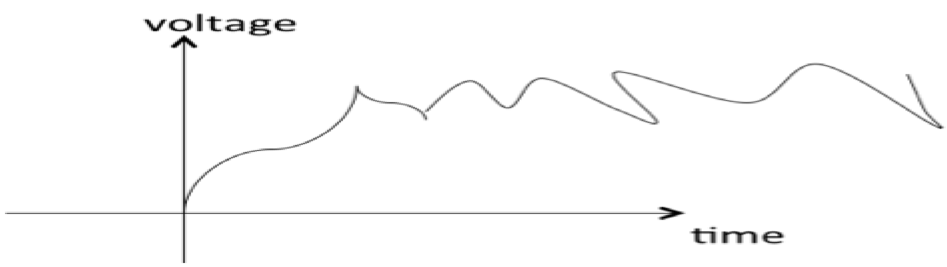


Figure3.10: Non-Deterministic Signal

iii. Even and Odd Signals

A signal is said to be even when it satisfies the condition x(t) = x(-t)

Example 1: t2, t4… cost etc.

Let x(t) = t2

x(-t) = (-t)2 = t2 = x(t)

∴, t2 is even function

Example 2: As shown in the following diagram, rectangle function x(t) = x(-t) so it is also even function.

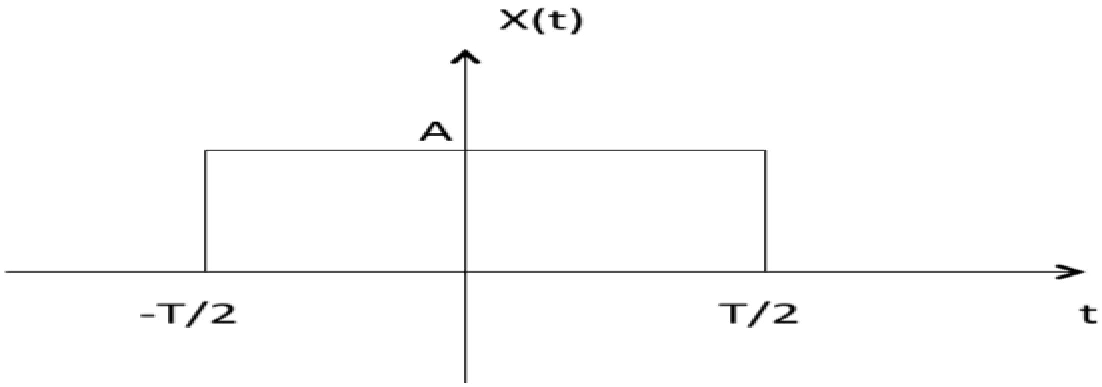


Figure3.11: Even and Odd Signals

A signal is said to be odd when it satisfies the condition x(t) = -x(-t)

Example 3: t, t3 ... And sin t

Let x(t) = sin t

x(-t) = sin(-t) = -sin t = -x(t)

∴, sin t is odd function.

Any function f(t) can be expressed as the sum of its even function $\boldsymbol{f_e(t)}$ and odd function $\boldsymbol{f_o(t)}$.

$\boldsymbol{f(t) = f_e(t) + f_0(t)}$

where

$\boldsymbol{f_e(t) = \frac{1}{2}[f(t) + f(-t)]}$

iv. Periodic and Aperiodic Signals

A signal is said to be periodic if it satisfies the condition

x(t) = x(t + T) or x(n) = x(n + N).

Where

T = fundamental time period,

1/T = f = fundamental frequency.

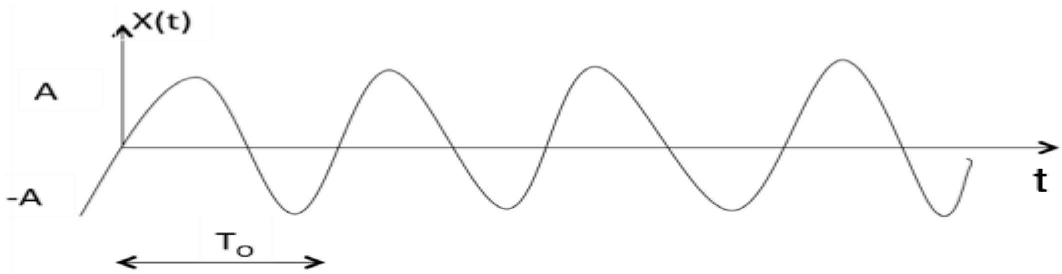


Figure3. 12: Periodic and aperiodic signals

The above signal will repeat for every time interval T0 hence it is periodic with period T0.

v. Energy and Power Signals

A signal is said to be energy signal when it has finite energy.

$$\textbf{Energy } E=\int_{-\infty}^{\infty} x^2(t)dt$$

A signal is said to be power signal when it has finite power.

$$\textbf{Power } P=\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} x^2(t)dt$$

NOTE: A signal cannot be both, energy and power simultaneously. In addition, a signal may be neither energy nor power signal.

Power of energy signal = 0

Energy of power signal = ∞

vi. Real and Imaginary Signals

A signal is said to be real when it satisfies the condition x(t) = x*(t)

A signal is said to be odd when it satisfies the condition x(t) = -x*(t)

Example:

If x(t)= 3 then x*(t)=3*=3 here x(t) is a real signal.

If x(t)= 3j then x*(t)=3j* = -3j = -x(t) hence x(t) is a odd signal.

Note: For a real signal, imaginary part should be zero. Similarly, for an imaginary signal, real part should be zero.

<u>Systems</u>

- Is a combination of elements, components that perform some task?
- is a set of element which produces o/p in response to i/p
- Mathematically, $y(n) = f[\ x(n)\ ]$
- Systems are classified into the following categories:
  - linear and Non-linear Systems
  - Time Variant and Time Invariant Systems
  - linear Time variant and linear Time invariant systems
  - Static and Dynamic Systems
  - Causal and Non-causal Systems
  - Invertible and Non-Invertible Systems
  - Stable and Unstable Systems

i. linear and Non-linear Systems

A system said to be linear when it satisfies superposition and homogenate principles. Consider two systems with inputs as $x_1(t)$, $x_2(t)$, and outputs as y_1 (t), y_2 (t) respectively. Then, according to the superposition and homogenate principles,

$$\mathbf{T\ [a_1\ x_1(t) + a_2\ x_2(t)] = a_1\ T[x_1(t)] + a_2\ T[x_2(t)]}$$

$$\mathbf{\therefore T\ [a_1\ x_1(t) + a_2\ x_2(t)] = a_1\ y_1\ (t) + a_2\ y_2\ (t)}$$

From the above expression, is clear that response of overall system is equal to response of individual system.

Example:

$$\mathbf{(t) = x^2(t)}$$

Solution:

$$\mathbf{y_1\ (t) = T[x_1(t)] = x_1^2(t)}$$

$$\mathbf{y_2\ (t) = T[x_2(t)] = x_2^2(t)}$$

$$\mathbf{T\ [a_1\ x_1(t) + a_2\ x_2(t)] = [a_1\ x_1(t) + a_2\ x_2(t)]^2}$$

Which is not equal to $\mathbf{a1\ y_1\ (t) + a_2\ y_2\ (t)}$. Hence, the system is said to be nonlinear.

ii. Time Variant and Time Invariant Systems

A system said to be time variant if its input and output characteristics vary with time. Otherwise, the system is consider as time invariant.

The condition for time invariant system is:

$$\mathbf{y\ (n\ ,\ t) = y(n-t)}$$

The condition for time variant system is:

$$\mathbf{y\ (n\ ,\ t) \neq y(n-t)}$$

Where $y(n, t) = T[x(n-t)]$ = input change

$y(n-t)$ = output change

Example:

$$\mathbf{y(n) = x(-n)}$$

$$\mathbf{y(n, t) = T[x(n-t)] = x(-n-t)}$$

$$\mathbf{y(n-t) = x(-(n-t)) = x(-n + t)}$$

$\therefore$ **$y(n, t) \neq y(n-t)$.** Hence, the system is time variant.

iii. linear Time variant (LTV) and linear Time Invariant (LTI) Systems

If a system is both linear and time variant, then it is called linear time variant (LTV) system. If a system is both linear and time Invariant then that system is called linear time invariant (LTI) system.

iv. Static and Dynamic Systems

Static system is memory-less whereas dynamic system is a memory system.

Example 1: **$y(t) = 2x(t)$**

For present value t=0, the system output is $y(0) = 2x(0)$. Here, the output is only dependent upon present input. Hence the system is memory less or static.

Example 2: **$y(t) = 2x(t) + 3x(t-3)$**

For present value t=0, the system output is $y(0) = 2x(0) + 3x(-3)$.

Here $x(-3)$ is past value for the present input for which the system requires memory to get this output. Hence, the system is a dynamic system.

v. Causal and Non-Causal Systems

A system is said to be causal if its output depends upon present and past inputs, and does not depend upon future input. For non causal system, the output depends upon future inputs also.

Example 1: **$y(n) = 2x(t) + 3x(t-3)$**

For present value t=1, the system output is $y(1) = 2x(1) + 3x(-2)$.

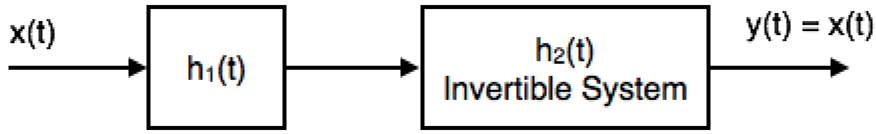
Here, the system output only depends upon present and past inputs. Hence, the system is causal.

Example 2: **$y(n) = 2x(t) + 3x(t-3) + 6x(t + 3)$**

For present value t=1, the system output is $y(1) = 2x(1) + 3x(-2) + 6x(4)$ Here, the system output depends upon future input. Hence the system is non-causal system.

vi. Invertible and Non-Invertible systems

A system said to invertible if the input of the system appears at the output.



Invertible system

$Y(S) = X(S)\ H1(S)\ H2(S)$

$= X(S)\ H1(S) \cdot 1/(H1(S))$ Since $H2(S) = 1/(\ H1(S)\ )$

$\therefore\ Y(S) = X(S)$

$\rightarrow y(t) = x(t)$

Hence, the system is invertible.

If **$y(t) \neq x(t)$,** then the system is said to be non-invertible.

vii. Stable and Unstable Systems

The system is said to be stable only when the output is bounded for bounded input. For a bounded input, if the output is unbounded in the system then it is said to be unstable.

Note: For a bounded signal, amplitude is finite.

Example 1: **$y\ (t) = x^2(t)$**

Let the input is u(t) (unit step bounded input) then the output y(t) = u2(t) = u(t) = bounded output.

Hence, the system is stable.

Example 2: **$y\ (t) = \int x(t)dt$**

Let the input is u (t) (unit step bounded input) then the output $y(t) = \int u(t)dt$ = ramp signal (unbounded because amplitude of ramp is not finite it goes to infinite when t → infinite).

Hence, the system is unstable.

<u>Channel Capacity Theorem (Shannon- Hartley theorem)</u>

Suppose a source sends r messages per second, and the entropy of a message is H bits per message. The information rate is R = r H bits/second.

One can intuitively reason that, for a given communication system, as the information rate increases the number of errors per second will also increase. Surprisingly, however, this is not the case.

Shannon’s theorem:

• A given communication system has a maximum rate of information C known as the channel capacity.

• If the information rate R is less than C, then one can approach arbitrarily small error probabilities by using intelligent coding techniques.

• To get lower error probabilities, the encoder has to work on longer blocks of signal data. This entails longer delays and higher computational requirements.

Thus, if $R \leq C$ then transmission may be accomplished without error in the presence of noise.

Unfortunately, Shannon’s theorem is not a constructive proof — it merely states that such a coding method exists. The proof can therefore not be used to develop a coding method that reaches the channel capacity. The negation of this theorem is also true: if $R > C$, then errors cannot be avoided regardless of the coding technique used.

Shannon-Hartley theorem

Consider a bandlimited Gaussian channel operating in the presence of additive Gaussian noise:

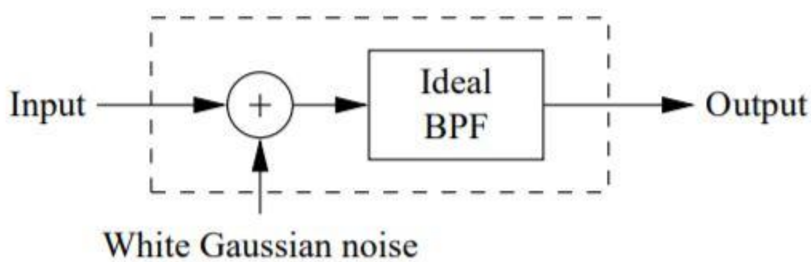


Figure 3.13: Shannon-Hartley Theorem

The Shannon-Hartley theorem states that the channel capacity given by

C = B log2 (1 + S/N)

Where C is the capacity in bits per second, B is the bandwidth of the channel in Hertz, and S/N is the signal-to-noise ratio.

****

Chapter 5

Overview of Data Communication Networking

A computer network is a system in which multiple computers connected to each other to share information and resources.

The term "**computer network**" is use to mean an interconnected collection of autonomous computer. Two computers are said to be interconnect only if they are able to exchange information.

Figure 5.1: Simply connected networks

Characteristics of a Computer Network

- Share resources from one computer to another.
- Create files and store them in one computer, access those files from the other computer(s) connected over the network.
- Connect a printer, scanner, or a fax machine to one computer within the network and let other computers of the network use the machines available over the network.

Following is the list of hardware's required to set up a computer network.

- Network Cables
- Distributors
- Routers
- Internal Network Cards
- External Network Cards

Types of Network:

1. **Personal Area Network**

A Personal Area Network (PAN) is smallest network, which is very personal to a user. This may include Bluetooth enabled devices or infrared enabled devices. PAN has connectivity range up to 10 meters. PAN may include wireless computer keyboard and mouse, Bluetooth enabled headphones, wireless printers and TV remotes. For example, Piconet is Bluetooth-enabled Personal Area Network, which may contain up to 8 devices connected together in a master-slave fashion.

Figure 5.2: PAN

2. Local Area Network

- A local area network (LAN) supplies networking capability to a group of computers in close proximity to each other such as in an office building, a school, or a home.
- A LAN is useful for sharing resources like files, printers, games or other applications.
- A LAN in turn often connects to other LANs, and to the Internet or other WAN.
- Most local area networks are built with relatively inexpensive hardware such as Ethernet cables, network adapters, and hubs. Wireless LAN and other more advanced LAN hardware options also exist.
- Specialized operating system software may be used to configure a local area network.
- For example, most flavors of Microsoft Windows provide a software package called Internet Connection Sharing (ICS) that supports controlled access to LAN resources.

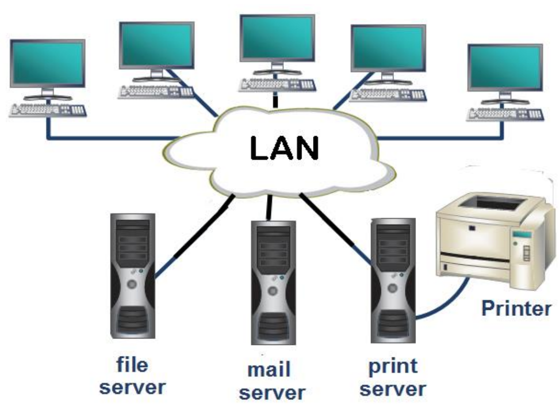


Figure 5.3: LAN

3. Wide Area Network

- A WAN spans a large geographic area, such as a state, province or country. WANs often connect multiple smaller networks, such as local area networks (LANs) or metro area networks (MANs).
- The world's most popular WAN is the Internet. Some segments of the Internet, like VPN-based extranets, are also WANs in themselves.

- Finally, many WANs are corporate or research networks that utilize leased lines.
- WANs generally utilize different and much more expensive networking equipment than do LANs. Key technologies often found in WANs include SONET, Frame Relay, and ATM.

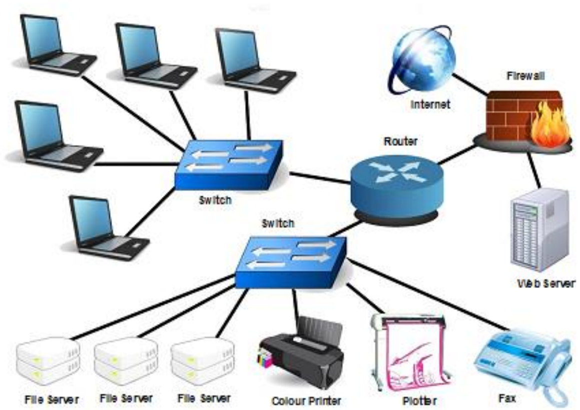


Figure 5.4: WAN

3. Metropolitan Area Network

- A Metropolitan Area Network (MAN) is one of a number of types of networks.
- A MAN is a relatively new class of network, it serves a role similar to an ISP, but for corporate users with large LANs. There are three important features which discriminate MANs from LANs or WANS:
  - The network size falls intermediate between LANs and WANs. A MAN typically covers an area of between 5 and 50 km diameter. Many MANs cover an area the size of a city.
  - A MAN (like a WAN) is not generally owned by a single organization. The MAN, its communications links and equipment are generally owned by either a consortium of users or by a single network provider who sells the service to the users. This level of service provided to each user must therefore can be negotiated with the MAN operator, and some performance guarantees are normally specified.
  - A MAN often acts as a high-speed network to allow sharing of regional re-sources (similar to a large LAN). It is also frequently used to provide a shared connection to other networks using a link to a WAN.

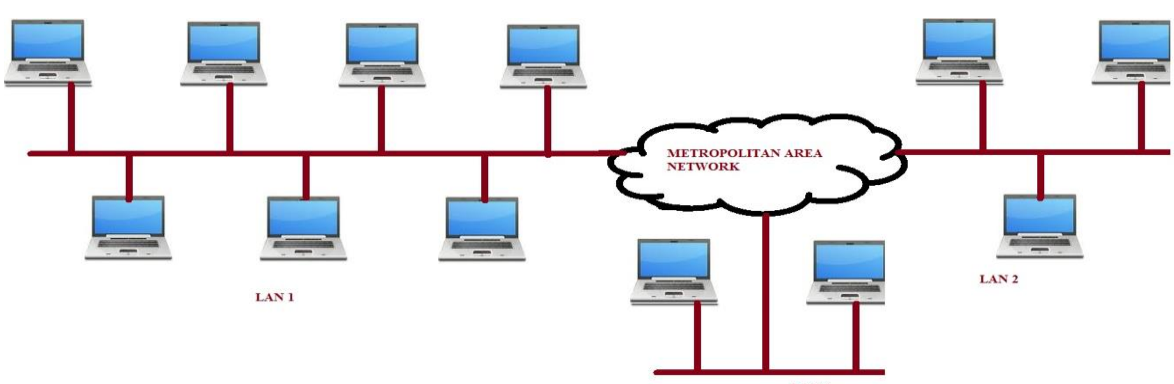


Figure 5.5: MAN

Network Topologies

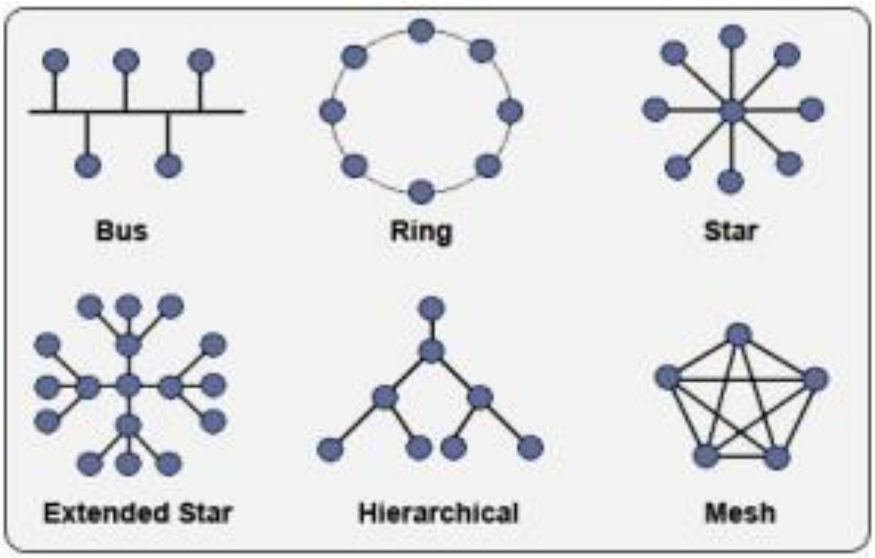


Figure 5.6: Network Topologies

i. ***Bus Topology***:

- In this type of network topology, all the nodes of a network are connected to a common transmission medium having two endpoints.
- All the data that travels over the network is transmitted through a common transmission medium known as the bus or the backbone of the network.
- When the transmission medium has exactly two endpoints, the network topology is known by the name, 'linear bus topology'. - In case the transmission medium, also called as the network backbone, has more than two endpoints, the network is said to have a distributed bus topology.
- Bus topology is easy to handle and implement and is best suited for small networks.
- But the downside of this topology is that the limited cable length limits the number of stations, thus limiting the performance to a less number of nodes.

ii. Ring Topology:

- In a ring topology, every node in the network is connected to two other nodes and the first and the last nodes are connected to each other.
- The data that are transmitted over the network pass through each of the nodes in the ring until they reach the destination node.
- In a ring network, the data and the signals that pass over the network travel in a single direction.

- The dual ring topology varies in having two connections between each of the network nodes.
- The data flow along two directions in the two rings formed thereby.
- The ring topology does not require a central server to manage connectivity be-tween the nodes and facilitates an orderly network operation.
- But, the failure of a single station in the network can render the entire network inoperable.
- Changes and moves in the stations forming the network affect the network operation.

iii. ***Mesh Topology:***

- In a full mesh network, each network node connected to every other node in the network.
- Due to this arrangement of nodes, it becomes possible for a simultaneous trans-mission of signals from one node to several other nodes.
- In a partially connected mesh network, only some of the network nodes connected to more than one node.
- This is beneficial over a fully connected mesh in terms of redundancy caused by the point-to-point links between all the nodes.
- The nodes of a mesh network require possessing some kind of routing logic so that the signals and the data traveling over the network take the shortest path during each of the transmissions.

iv. ***Star Topology:***

- In this type of network topology, each node of the network connected to a central node, which known is as a hub.
- The data that is transmitted between the network nodes passes across the central hub.
- A distributed star formed by the interconnection of two or more individual star networks.
- The centralized nature of a star network provides a certain amount of simplicity while also achieving isolation of each device in the network.
- However, the disadvantage of a star topology is that the network transmission is largely dependent on the central hub. - The failure of the central hub results in total network inoperability.

v. ***Tree Topology:***

- It is known's as a hierarchical topology and has a central root node that connected to one or more nodes of a lower hierarchy.
- In a symmetrical hierarchy, each node in the network has a specific fixed number of nodes connected to those at a lower level.
- Apart from these basic types of network topologies, there are hybrid network topologies, which are composed of a combination of two or more basic topologies.
- These network mappings aim at harnessing the advantages of each of the basic topologies used in them.

- Network topologies are the physical arrangements of network nodes and wires. What is interesting is that the inanimate nodes and wires turn 'live' for the trans-mission of information!

Protocol Architecture:

Is the layered structure of hardware and software that supports the exchange of data between systems and supports applications such a as electronic mail and file transfer.

The key features of protocol are:

- Syntax: concerns the format of the data blocks
- Semantics: Includes control information for coordination and error handling
- Timing: Includes speed matching and sequencing

Open System Interconnection (OSI)

The Open System Interconnection (OSI) model defines a networking framework to implement protocols in seven layers. Use this handy guide to compare the different layers of the OSI model and understand how they interact with each other. The OSI model does not perform any functions in the networking process. It is a conceptual framework so we can better understand the complex interactions that are happening.

The International Standards Organization (ISO) developed the Open Systems Interconnection (OSI) model. It divides network communication into seven layers. In this model, layers 1-4 are consider the lower layers, and mostly concern themselves with moving data around. Layers 5-7, called the upper layers, contain application-level data. Networks operate on one basic principle: "pass it on." Each layer takes care of a very specific job, and then passes the data onto the next layer.

In the OSI model, control is pass from one layer to the next, starting at the application layer (Layer 7) in one station, and proceeding to the bottom layer, over the channel to the next station and back up the hierarchy. The OSI model takes the task of inter-networking and divides that up into what referred to as a vertical stack that consists of the following 7 layers.

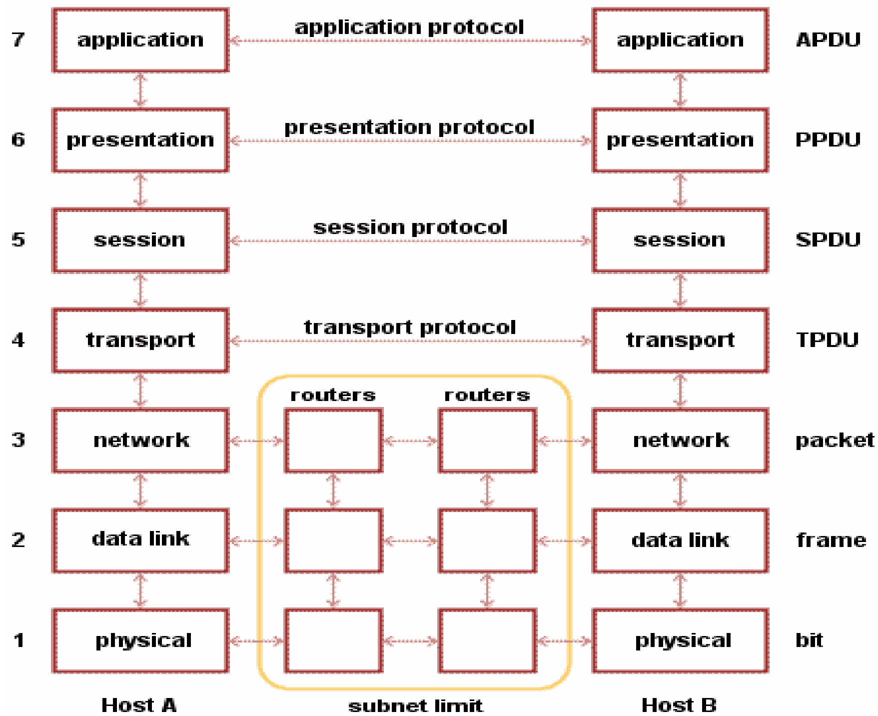


Figure 5.7: OSI Layer

1. Physical Layer:

- A physical layer covers the physical interface between devices and the rules by which bits passed from one to another.
- It relates to the physical properties of the interface to a transmission medium.
- For example, connector that joins one or more circuits.
- Electrical part of physical layer relates to the representation of bits.
- Functional parts of physical layer specify the function performed by individual circuits between a system and the transmission medium.
- Similarly, procedural part of physical layer species the sequence of events by which bit streams are exchange across the physical medium.

2. Data Link Layer:

- Data link layer attempts to make the physical link reliable and provides the means to activate, maintain, and deactivate the link.
- It provides for the reliable transfer of information across physical link.
- It sends blocks with the necessary synchronization, error control and flow control.

3. Network Layer:

- Determines how data are transferred between network devices
- Routes packets according to unique network device addresses
- Provides flow and congestion control to prevent network resource depletion

4. Transport Layer:

- It provides the mechanism for the exchange of data between and system.
- The connection oriented transport service ensures that data are deliver error free, in sequence with no loss or duplication.
- Manages end-to-end message delivery in network
- Provides reliable and sequential packet delivery through error recovery and flow control mechanisms
- Provides connectionless oriented packet delivery

5. Session Layer:

- It provides the mechanism for controlling the dialog between application in and systems.
- Manages user sessions and dialogues
- Controls establishment and termination of logic links between users
- Reports upper layer errors

6. Presentation Layer:

- It defines the format of the date to be exchange between applications.
- It defines the syntax used between applications and provides for the selection and subsequent modification of the presentation used.
- Masks the differences of data formats between dissimilar systems
- Specifies architecture-independent data transfer format
- Encodes and decodes data; encrypts and decrypts data; compresses and decompresses data

7. Application Layer:

- Defines interface-to-user processes for communication and data transfer in net-work
- Provides standardized services such as virtual terminal, file and job transfer and operations.

TCP/IP Protocol Architecture

The OSI Model we just looked at is just a reference/logical model. It was design to describe the functions of the communication system by dividing the communication procedure into smaller and simpler components. However, when we talk about the TCP/IP model, it was designed and developed by Department of Defense (DOD) in 1960s and is based on standard protocols. It stands for Transmission Control Protocol/Internet Protocol. The TCP/IP model is a concise version of the OSI model. It contains four layers, unlike seven layers in the OSI model. The layers are:

1. Process/Application Layer
2. Host-to-Host/Transport Layer
3. Internet Layer
4. Network Access/Link Layer

TCP/IP MODEL
Application Layer
Transport Layer
Internet Layer
Network Access Layer

OSI MODEL
Application Layer
Presentation Layer
Session Layer
Transport Layer
Network Layer
Data Link Layer
Physical Layer

Figure5.8: Diagrammatic comparison of TCP/IP and OSI model

TCP/IP functionality is divided into four layers, each of which include specific protocols:

- **The application layer** provides applications with standardized data exchange. Its protocols include the HTTP, FTP, Post Office Protocol 3 (POP3), Simple Mail Transfer Protocol (SMTP) and Simple Network Management Protocol (SNMP). At the application layer, the payload is the actual application data.

- **The transport layer** is responsible for maintaining end-to-end communications across the network. TCP handles communications between hosts and provides flow control, multiplexing and reliability. The transport protocols include TCP and User Datagram Protocol (UDP), which is sometimes used instead of TCP for special purposes.

- **The network layer** also called the internet layer, deals with packets and connects independent networks to transport the packets across network boundaries. The network layer protocols are the IP and the Internet Control Message Protocol (ICMP), which is used for error reporting.

- **The physical layer**, also known as the network interface layer or data link layer, consists of protocols that operate only on a link -- the network component that interconnects nodes or hosts in the network. The protocols in this lowest layer include Ethernet for local area networks (LANs) and the Address Resolution Protocol (ARP).

Difference between TCP/IP and OSI

- TCP refers to Transmission Control Protocol where as OSI refers to Open Systems Interconnection.
- TCP/IP has four layers & OSI has seven layers.
- TCP/IP is more reliable than OSI is less reliable

- TCP/IP does not have very strict boundaries where as OSI has strict boundaries
- TCP/IP follow a horizontal approach where as OSI follows a vertical approach.
- TCP/IP uses both session and presentation layer in the application layer itself and OSI uses different session and presentation layers.
- TCP/IP developed protocols then model where as OSI developed model then protocol.
- Transport layer in TCP/IP does not provide assurance delivery of packets and in OSI; model transport layer provides assurance delivery of packets.
- TCP/IP model network layer only provides connection less services whereas Connection less and connection oriented both services are provided by network layer in OSI model.
- Protocols cannot be replaced easily in TCP/IP model While in OSI model; Protocols are better covered and is easy to replace with the change in technology.

<u>Local Area Network Architectures, LLC/MAC & Routing</u>

Local Area Network (LAN) is a data communication network connecting various terminals or computers within a building or limited geographical area. The connection among the devices could be wired or wireless. Ethernet, Token Ring and Wireless LAN using IEEE 802.11 are examples of standard LAN technologies.

LLC / MAC

The Open System Interconnections (OSI) model is a 7 – layered networking framework that conceptualizes how communications should be done between heterogeneous systems. The data link layer is the second lowest layer. It is divided into two sublayers −

- The logical link control (LLC) sublayer
- The medium access control (MAC) sublayer

The following diagram depicts the position of the LLC sublayer –

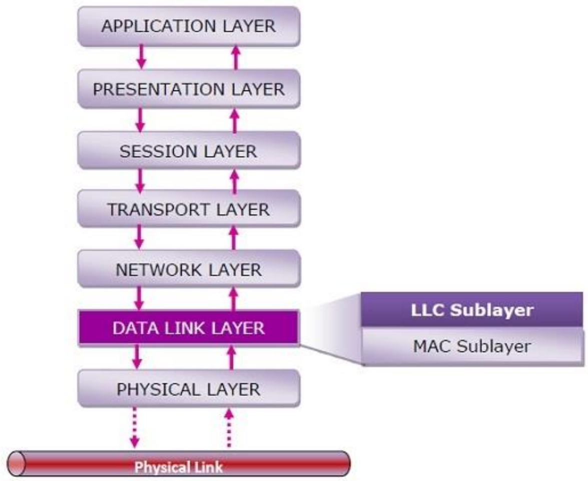


Figure 5.9: LLC and MAC in OSI Layer

Functions of LLC Sublayer

- The primary function of LLC is to multiplex protocols over the MAC layer while transmitting and likewise to de-multiplex the protocols while receiving.
- LLC provides hop-to-hop flow and error control.
- It allows multipoint communication over computer network.
- Frame Sequence Numbers are assign by LLC.
- In case of acknowledged services, it tracks acknowledgements

Function of MAC sublayer

- It provides an abstraction of the physical layer to the LLC and upper layers of the OSI network.
- It is responsible for encapsulating frames so that they are suitable for transmission via the physical medium.
- It resolves the addressing of source station as well as the destination station, or groups of destination stations.
- It performs multiple access resolutions when more than one data frame to be transmit. It determines the channel access methods for transmission.
- It also performs collision resolution and initiating retransmission in case of collisions.
- It generates the frame check sequences and thus contributes to protection against transmission errors.

Ethernet (IEEE 802.3)

Ethernet is most widely used LAN Technology, which is define under IEEE standards 802.3. The reason behind its wide usability is Ethernet is easy to understand, implement, maintain and allows low-cost network implementation. In addition, Ethernet offers flexibility in terms of topologies, which are allowed. Ethernet generally uses Bus Topology. Ethernet operates in two layers of the OSI model, Physical Layer, and Data Link Layer. For Ethernet, the protocol data unit is Frame since we mainly deal with DLL. In order to handle collision, the Access control mechanism used in Ethernet is CSMA/CD.

Currently, these data rates are define for operation over optical fibers and twisted-pair cables:

i. **Fast Ethernet:-** Fast Ethernet refers to an Ethernet network that can transfer data at a rate of 100 Mbit/s.
ii. **Gigabit Ethernet:-** Gigabit Ethernet delivers a data rate of 1,000 Mbit/s (1 Gbit/s).
iii. **10 Gigabit Ethernet:-** 10 Gigabit Ethernet is the recent generation and delivers a data rate of 10 Gbit/s (10,000 Mbit/s). It is generally use for backbones in high-end applications requiring high data rates.

Aloha

The Aloha protocol was design as part of a project at the University of Hawaii. It provided data transmission between computers on several of the Hawaiian Islands involving packet radio networks. Aloha, is a multiple access protocol at the data link layer and proposes how multiple terminals access the medium without interference or collision.

There are two different versions of ALOHA:

1. Pure Aloha

Pure Aloha is an un-slotted, decentralized, and simple to implement a protocol. In pure ALOHA, the stations simply transmit frames whenever they want data to send. It does not check whether the channel is busy or not before transmitting. In case, two or more stations transmit simultaneously, the collision occurs and frames are destroyed. Whenever any station transmits a frame, it expects the acknowledgment from the receiver. If it is not received within a specified time, the station assumes that the frame or acknowledgment has been destroyed. Then, the station waits for a random amount of time and sends the frame again. This randomness helps in avoiding more collisions. This scheme works well in small networks where the load is not much. But in largely loaded networks, this scheme fails poorly. This led to the development of Slotted Aloha.

To assure pure aloha: Its throughput and rate of transmission of the frame to be predicted.

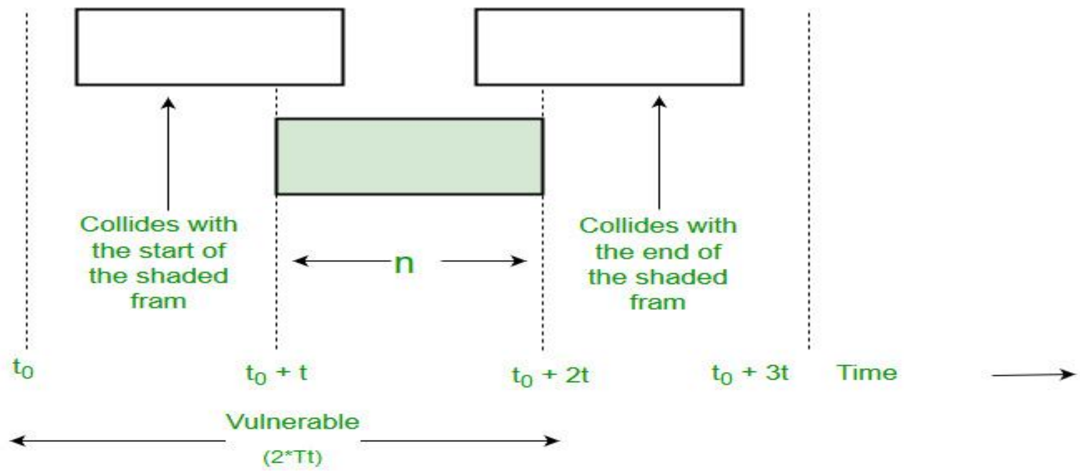


Figure 5.10: Pure Aloha

For that to make some assumption:

i) All the frames should be the same length.
ii) Stations cannot generate frame while transmitting or trying to transmit frame.
iii) The population of stations attempts to transmit (both new frames and old frames that collided) according to a Poisson distribution.

Vulnerable Time = 2 * Tt

Efficiency of Pure ALOHA:

Spure= G * e^-2G

where G is number of stations wants to transmit in Tt slot.

Maximum Efficiency:

Maximum Efficiency will be obtained when G=1/2

(Spure)max = 1/2 * e^-1

= 0.184

Which means, in **Pure ALOHA**, only about **18.4%** of the time used for successful transmissions.

2. Slotted Aloha

This is quite similar to Pure Aloha, differing only in the way transmissions take place. Instead of transmitting right at demand time, the sender waits for some time. In slotted ALOHA, the time of the shared channel divided into discrete intervals called Slots. The stations are eligible to send a frame only at the beginning of the slot and only one frame per slot is send. If any station is not able to place the frame onto the channel at the beginning of the slot, it has to wait until the beginning of the next time slot. There is still a possibility of collision if two stations try to send at the beginning of the same time slot. Still the number of collisions that can possibly take place is reduce by a large margin and the performance becomes much well compared to Pure Aloha.

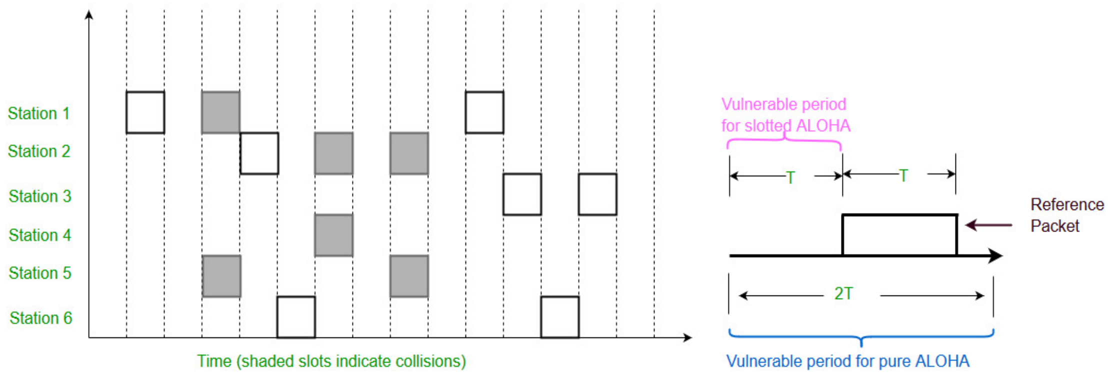


Figure 5.11: Slotted Aloha

Collision is possible for only the current slot. Therefore, Vulnerable Time is Tt.

Efficiency of Slotted ALOHA:

Sslotted = G * e^-G

Maximum Efficiency:

(Sslotted)max = 1 * e^-1

= 1/e = 0.368

Maximum Efficiency, in Slotted ALOHA, is 36.8%.

Carrier Sense Multiple Access (CSMA)

This method was develop to decrease the chances of collisions when two or more stations start sending their signals over the datalink layer. Carrier Sense multiple access requires that each station first check the state of the medium before sending.

Vulnerable time = Propagation time (Tp)

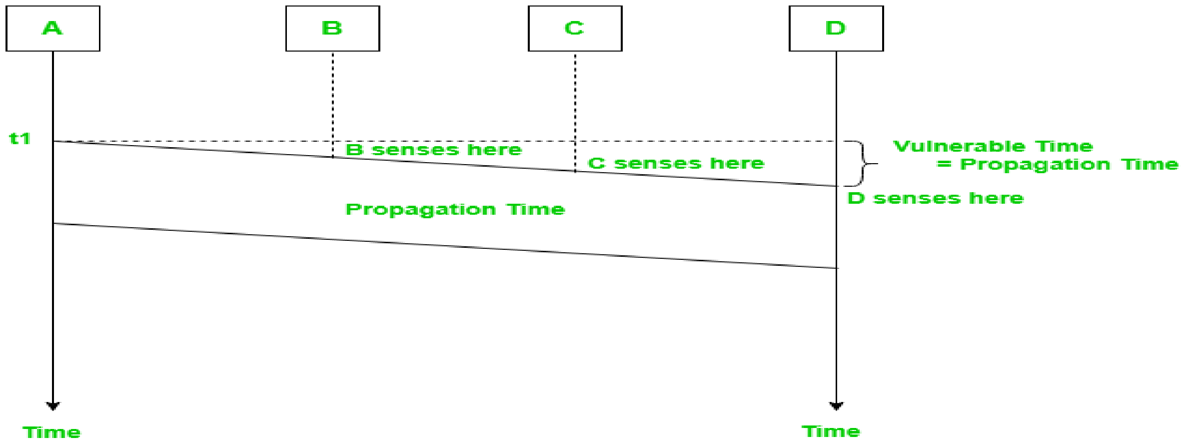


Figure 5.142: CSMA

The persistence methods can be applied to help the station take action when the channel is busy/idle.

1. Carrier Sense Multiple Access with Collision Detection (CSMA/CD) –

In this method, a station monitors the medium after it sends a frame to see if the transmission was successful. If successful, the station is finished, if not, the frame is sent again.

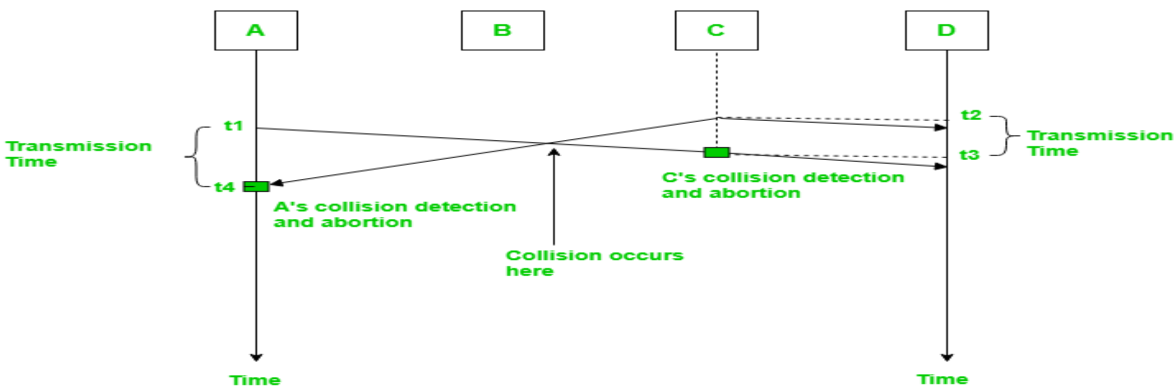


Figure 5.13: CSMA/ CD

In the diagram, A starts send the first bit of its frame at t1 and since C sees the channel idle at t2, starts sending its frame at t2. C detects A's frame at t3 and aborts transmission. A detects C's frame at t4 and aborts its transmission. Transmission time for C's frame is therefore t3-t2 and for A's frame is t4-t1.

So, the frame transmission time (Tfr) should be at least twice the maximum propagation time (Tp). This can be deduced when the two stations involved in collision are maximum distance apart.

Process –

The entire process of collision detection can be explain as follows:

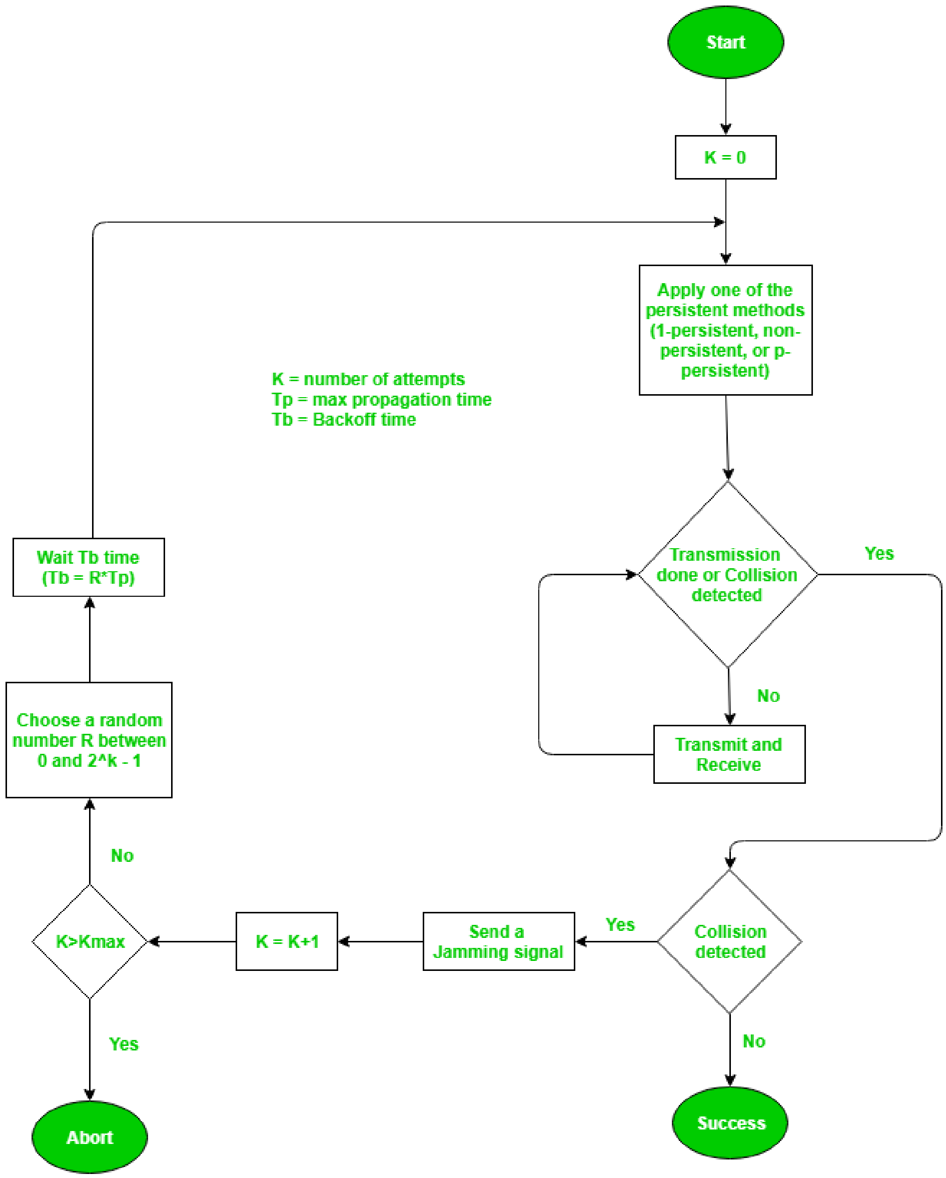


Figure 5.14: Algorithm for CSMA/CD

Throughput and Efficiency – The throughput of CSMA/CD is much greater than pure or slotted ALOHA.

- For 1-persistent method throughput is 50% when G=1.
- For non-persistent method throughput can go up to 90%.

2. Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) –

The basic idea behind CSMA/CA is that the station should be able to receive while transmitting to detect a collision from different stations. In wired networks, if a collision has occurred then the energy of received signal almost doubles and the station can sense the possibility of collision. In case of wireless networks, most of the energy is used for transmission and the energy of received signal increases by only

5-10% if collision occurs. It can’t be used by station to sense collision. Therefore, CSMA/CA has been specially design for wireless networks.

These are three type of strategies:

- InterFrame Space (IFS) – When a station finds the channel busy, it waits for a period called IFS time. IFS can also be used to define the priority of a station or a frame. Higher the IFS lower is the priority.
- Contention Window – It is the amount of time divided into slots. A station, which is ready to send frames, chooses random number of slots as wait time.
- Acknowledgements – The positive acknowledgements and time-out timer can help guarantee a successful transmission of the frame.

The entire process for collision avoidance can be explain as follows:

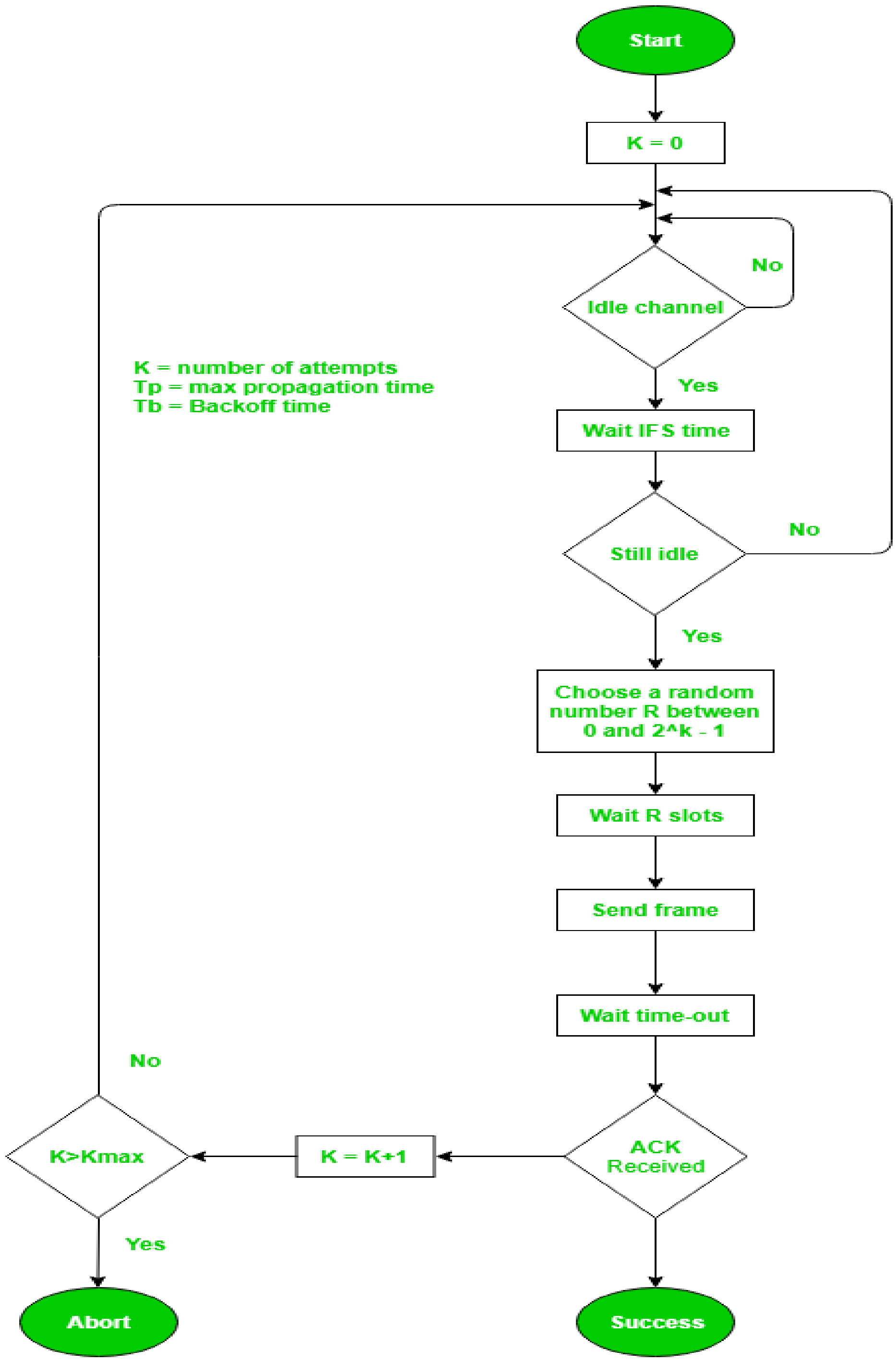


Figure 5.15: Algorithm for CSMA/ CA

5.3 Wide Area Networks

X.25

X.25 is a protocol suite defined by ITU-T for packet switched communications over WAN (Wide Area Network). It was originally design for use in the 1970s and became very popular in 1980s. Presently, it is use for networks for ATMs and credit card verification. It allows multiple logical channels to use the same physical line. It also permits data exchange between terminals with different communication speeds.

X.25 has three protocol layers

- Physical Layer: It lays out the physical, electrical and functional characteristics that interface between the computer terminal and the link to the packet switched node. X.21 physical implementer is commonly use for the linking.
- Data Link Layer: It comprises the link access procedures for exchanging data over the link. Here, control information for transmission over the link is attach to the packets from the packet layer to form the LAPB frame (Link Access Procedure Balanced). This service ensures a bit-oriented, error-free, and ordered delivery of frames.
- Packet Layer: This layer defines the format of data packets and the procedures for control and transmission of the data packets. It provides external virtual circuit service. Virtual circuits may be of two types: virtual call and permanent virtual circuit. The virtual call is establish dynamically when needed through call set up procedure, and the circuit is relinquished through call clearing procedure. Permanent virtual circuit, on the other hand, is fixed and network assigned.

Equipment used

- X.21 implementer
- DTE : Data Terminal Equipment
- DCTE: Data Circuit Terminating Equipment

Frame Relay

Frame Relay is a packet switched communication service from LANs (Local Area Network) to backbone networks and WANs. It operates at two layers: physical layer and data link layer. It supports all standard physical layer protocols. It is mostly implement at the data link layer.

Frame Relay uses virtual circuits to connect a single router to multiple remote sites. In most cases, permanent virtual circuits are used, i.e. a fixed network-assigned circuit is use through which the user sees a continuous uninterrupted line. However, switched virtual circuits may also be used.

Frame relay is a fast packet technology based on X.25. Data is transmit by encapsulating them in multiple sized frames. The protocol does not attempt to correct errors and so it is faster. Error correction is handle by the endpoints, which are responsible for retransmission of dropped frames.

Frame Relay Devices are

- DTE: Data Terminal Equipment
- DCTE: Data Circuit Terminating Equipment

ATM

- Asynchronous Transfer Mode
- It is a streamlined packet transfer interface.

- ATM makes use of fixed size packets called cells.
- The use of fixed size and fixed formats results an efficient scheme for transmission over high-speed networks.
- data rate ranges from 25.6 Mbps to 622.08 Mbps
- Physical layer specifies transmission medium and signal encoding scheme
- ATM layer defines transmission of data in fixed size cells and defines the use of logical connection.
- ATM adaptation layer maps higher layer information into ATM cells to be trans-ported over an ATM network.
- User plane provides user information into ATM cells to be transported over an ATM network
- user plane provides user information transfer (eg. flow control, error control)
- Control plane provides call control and connection control functions.
- Management plane performs coordination between all the planes and layers management.

****

Chapter 6

Transmission Media

Introduction

Transmission media is a communication channel that carries the information from the sender to the receiver. Data transmit through the electromagnetic signals. The main functionality of the transmission media is to carry the information in the form of bits through LAN (Local Area Network). It is a physical path between transmitter and receiver in data communication. In a copper-based network, the bits in the form of electrical signals. In a fiber based network, the bits in the form of light pulses. In OSI (Open System Interconnection) phase, transmission media supports the Layer 1. Therefore, it is consider being as a Layer 1 component. The electrical signals can be send through the copper wire, fiber optics, atmosphere, water, and vacuum. The characteristics and quality of data transmission are determine by the characteristics of medium and signal. Transmission media is of two types are wired media and wireless media. In wired media, medium characteristics are more important whereas, in wireless media, signal characteristics are more important. Different transmission media have different properties such as bandwidth, delay, cost and ease of installation and maintenance. The transmission media is available in the lowest layer of the OSI reference model, i.e., Physical layer.

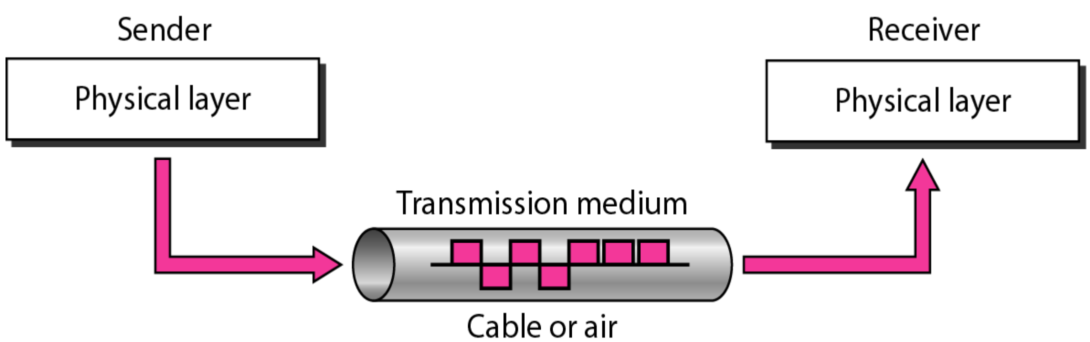


Figure 1: Transmission medium connection

Types of Transmission media

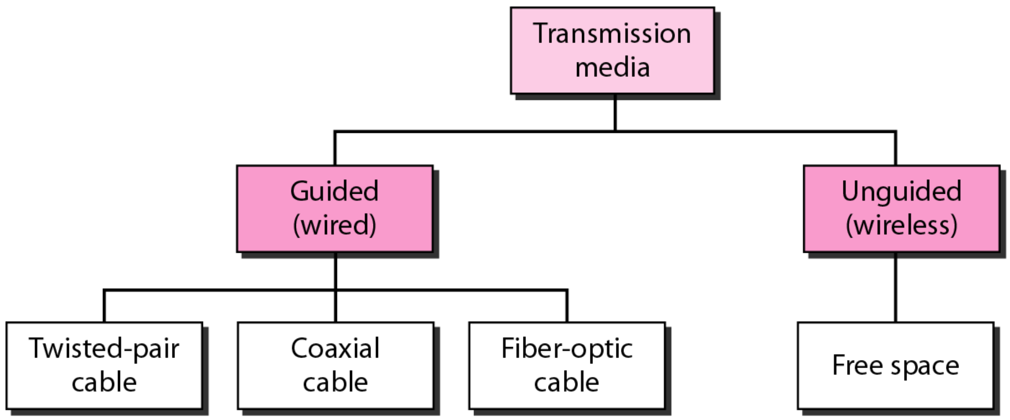


Figure 15: Types of Transmission media

Guided and Unguided Media:

- **Guided media** are those that provides physical conduction from one device to another which includes **twisted pairs, co-axial cables and fiber-optic cables.**
- **Unguided Media** transports electromagnetic waves without using a physical conductor.
    - This type of communication is often refer to as wireless communication.

Transmission Media:

- Transmission medium is the physical path between transmitter and in a data transmission system. - Transmission media can be classified or unguided.
- In both cases, communication is in the form of electromagnetic waves.
- With guided media, the waves are guided along a solid medium, such as copper twisted pair, copper coaxial cable, and optical fiber.
- The atmosphere and outer space are examples of unguided media that provide a means of transmitting
- Electromagnetic signals but do not guide them; this form of transmission is usually referred to as wireless transmission.
- The characteristics and quality of a data transmission are determine by both the characteristics of the medium and the characteristics of the signal.
- In the case of guided media, the medium itself is more important in determining the limitations of transmission.
- For unguided media, the bandwidth of the signal produced by the transmitting antenna is more important than the medium in determining transmission characteristics.
- One key property of signals transmitted by antenna is directionality.
- In general, signals at lower frequencies are omnidirectional, i.e. the signal propagates in all directions from the antenna.
- At higher frequencies, it is possible to focus the signal into a directional beam.
- In considering the design of data transmission systems, a key concern, generally, is data rate and distance: the greater the data rate and distance, the better.
- A number of design factors relating to the transmission medium and to the signal determine the data rate and distance:

Bandwidth:

- All other factors remaining constant, the greater the bandwidth of a signal, the higher the data rate that can be achieve.

Transmission impairments:

- Impairments, such as attenuation, limit the distance.
- For guided media, twisted pair generally suffer more impairment than coaxial cable, which in turn suffers more than optical fiber.

Interference:

- Interference from competing signals in overlapping frequency bands can distort or wipe out a signal. Interference is of particular concern for unguided media, but it is also a problem with guided media.
- For guided media, interference can be cause by emanations from nearby cables.
- For example, twisted pair are often bundled together, and conduits often carry multiple cables. - Interference can also be experience from unguided transmissions.
- Proper shielding of a guided medium can minimize this problem.
- Number of receivers:
  - A guided medium can be used to construct a point-to point link or a shared link with multiple attachments.
  - In the latter case, each attachment introduces some attenuation and distortion on the line, limiting distance and/or data rate.

Wired Pairs:

- Wires are described by their size.
- Higher gauge number indicates thinner wire size.
- The smaller the diameter of the wire, the greater is resistance to the propagation of a signal.
- Increased resistance results in a decreased bit rate across the communication path.
- At higher transmission frequencies, the signal tends to travel on the outside surface of the wire.
- A small wire provides less total surface for the radiating signal, resulting in increased signal loss.
- The local subscriber loops (of the telephone system) and usually to 22-26-gauge wire.
- Trunk and toll lines typically employ 19-gauge wires.
- Several Hundred of these wires are packaged into one cables.
- The wires are paired and twisted around each other to decrease certain electromagnetic problems.
- The most common twisted pair cable used in communications is referred to as unshielded twisted pair(UTP) cable.
- STP cable (shielded Twisted Pair) has a metal foil or braided-mesh covering each pair of insulated conductors.
- Although metal casing improves the quality of cable by preventing the penetration of noise or cross-talk, it is bulkier and more expensive.

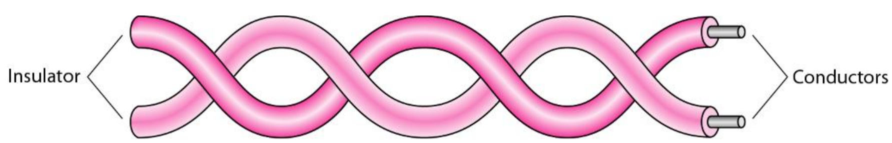


Figure 3:Twisted Pair Cables

Applications:

- Twisted pair cables are used in telephone lines to provide voice and data channels.
- The line that connects subscribers to the central telephone office is most commonly unshielded twisted pair cables.
- Local area networks such as 10 Base-T and 100 Base-T also use twisted pair cables.

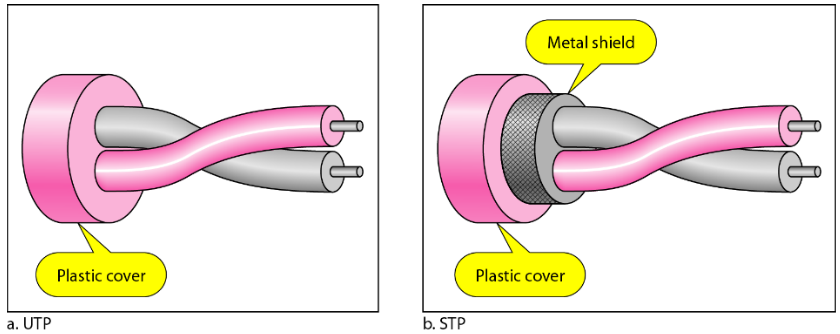


Figure 4: UTP and STP Cables

Microwaves:

- Microwave is a directed line of sight (LOS) radio transmission.
- It is used for wide band communication systems and is quite common in the telephone system.
- Television transmission also utilizes microwave transmission because microwave transmission is above the 1 GHz and provides the capacity required for video transmission.
- The high bandwidth gives small wavelength and the smaller the wavelength, the smaller one can design the microwave antenna.
- The antenna size has significant implications for distributed processing systems.
- The transmitting towers are spaced 20-30 m apart.
- Transmitted radio bean is focused to the receiving antenna.

COAXIAL Cables:

- Co-axial cables carry signals of higher frequency ranges than twisted pair cable.
- Instead of having two wires, co-axial cable has a central core conductor of solid or standard wire (usually copper) enclosed in an insulating sheath which in turn is encased in an outer conductor of metal foil or combination of two.
- The outer metallic wrapping serves both as a shield against noise and as a second conductor.
- The whole cable is protected by a plastic cover.

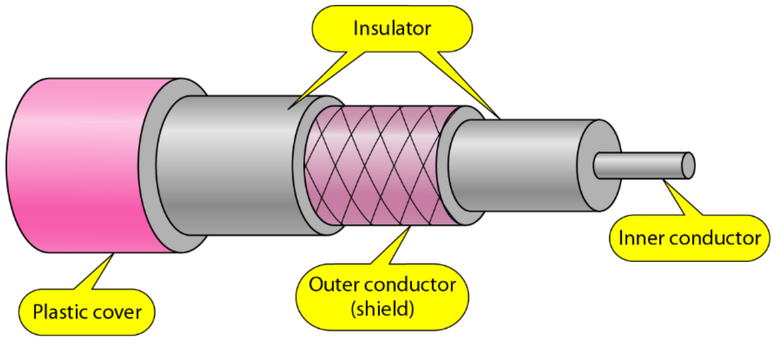


Figure 5: Coaxial Cables

Applications:

- The use of co-axial cable is diverse but nowadays it is shrinking due to fiber optic cable.
- Co-axial cables are used in analog telephone networks and cable Tv networks.

Fiber optic Cables:

- A fiber optic cable is made of glass or plastic and transmits signals in the form of lights.
- Lights travels in a straight line as long as it is moving through a single uniform substance.
- If a ray of light travelling through one substance suddenly enters another (more or less dense the ray changes direction).

- As the above figure, if the angle of incidence is less than the critical angle, the ray diffracts and move to closer to the surface.
- If angle of incidence is equal to the critical angle, the light bends along the interface and refraction occurs.
- If the angle is greater than the critical angle, the ray reflects and travels again in the denser substances.

- Optical fibers use reflection to guide light through a channel.
- A glass or plastic core is surrounded by cladding of less dense glass or plastic.
- The difference in density of the two materials must be such that a beam of light moving through the core is reflected off the cladding instead of being reflected into it.

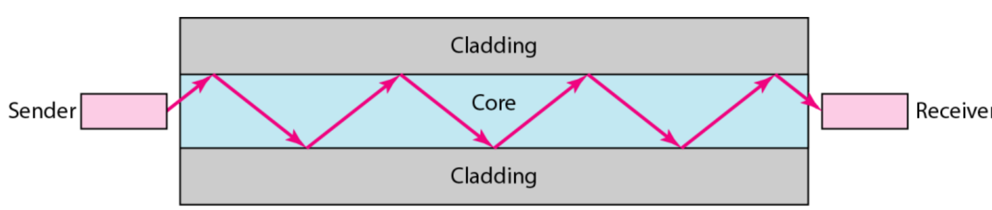


Figure 6: Optical Fiber

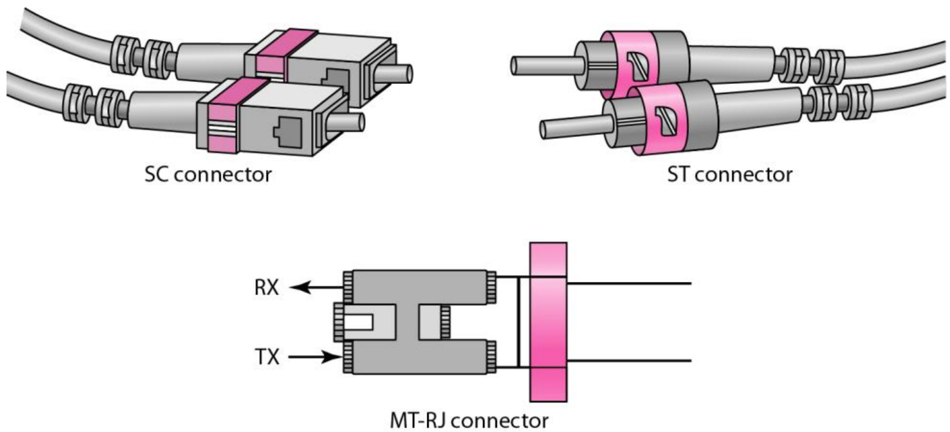


Figure 7: Fiber optic connector

Propagation Modes:

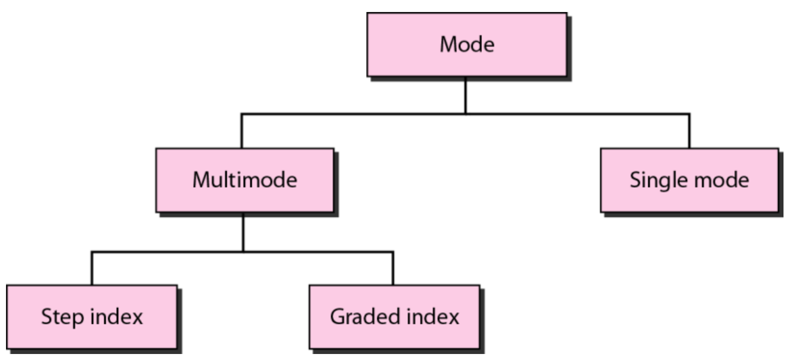


Figure 8: Types of Propagation mode

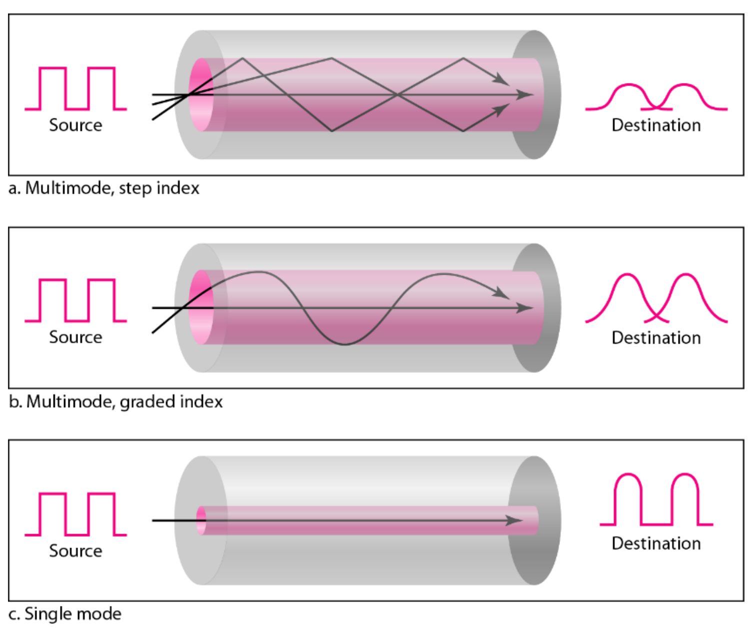


Figure 9: Modes of propagation

Multimode step index:

- in multimode step-index fiber, the density of the core remains constant from the center of the edges,
- A beam of light moves through these constant densities in a straight line until it reaches the interface of the core and the cladding.
- As the interface, there is an abrupt change to a lower density that alters the angle of the beam motion.
- The term step index refers to the suddenness of this change.

Multimode graded index:

- A second type of fiber called multimode graded index fiber is one with varying densities.
- Density is highest at center of the core and decreases gradually to its lowest at the edge.

Single mode:

- Single mode uses step index fiber and a highly focused source of light that limits beams to a small range of angles, all close to the horizontal.
- The single mode fiber is manufactured with a much smaller diameter that that of multimode fiber.

Applications:

- Optical fiber cable is found in backbone networks because of its wide bandwidth and is cost effective.
- cable TV companies use a combination of optical fibers and co-axial cable, thus creating a hybrid network.

Advantages:

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

Disadvantages:

- Installation/maintenance
- Unidirectional
- Cost

Electromagnetic waves:

- The electromagnetic waves used e propagation characteristics of wireless channels are highly dependent on frequency.
- Since, electromagnetic waves don't need any medium to transmit the signal specially in wireless communication system we often use the atmosphere for transmission of channel.
- Here, interference and propagation condition are strongly dependent upon the frequency.

Types of Electromagnetic waves:

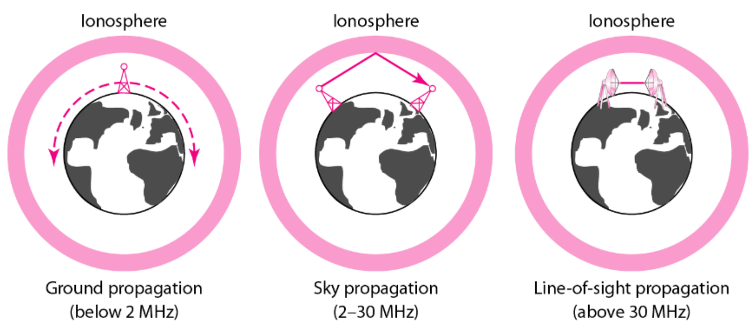


Figure 10: Propagation Method

Ground Wave Propagation:

- Dominant mode of propagation for frequency below 2 MHz
- Electromagnetic waves are guided by the conducting surface of the earth, along which they are propagated.
- Diffraction of the wave causes it to propagate where this propagation mode is used in AM broadcasting.
- For efficient radiation, the antenna needs to be longer than 1/10 th of the wave length.

Sky Wave Propagation:

- Dominant mode of propagation for frequencies in between 2 to 30 Mhz.
- Long distance coverage is obtained by the reflecting the wave at the ionosphere and at the earth boundaries.
- This is caused due to reflection.

Line of sight(LOS) or Space Wave propagation:

- Dominant mode of propagation for frequencies above 30 Mhz.
- Here, electromagnetic wave propagates in a straight line.
- Very little reflection by the ionosphere,
- This is used for satellite communication.
- Its maximum range is limited to line of sight due to nature of propagation.

Disadvantages:

- For communication between two each stations, the signal path has to be above the horizon otherwise they will block the LOS path.
- Thus antennas need to be placed on tall towers so that receiver antenna can see the transmitting antenna.

Wireless transmission

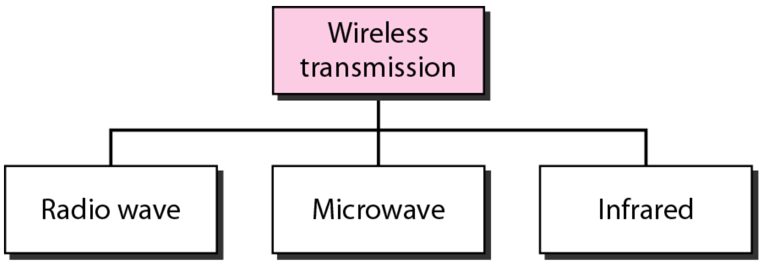


Figure 11: types of wireless transmission

Radio waves and Microwaves:

- 3 KHz to 1 GHz
- 1 GHz to 300 GHz
- for the most part is omnidirectional
- unidirectional
- Omnidirectional antennas are generally used
- line of sight propagation, unidirectional antennas are used
- AM, FM radio
- cellular phones, satellite networks

Satellite Communications:

- Satellite communications are comprised of 2 main components:

The Satellite

- The satellite itself is also known as the space segment, and is composed of three separate units, namely the fuel system, the satellite and telemetry controls, and the transponder.
- The transponder includes the receiving antenna to pick-up signals from the ground station, a broad band receiver, an input multiplexer, and a frequency converter which is used to reroute the received signals through a high-powered amplifier for downlink.
- The primary role of a satellite is to reflect electronic signals.
- In the case of a telecom satellite, the primary task is to receive signals from a ground station and send them down to another ground station located a considerable distance away from the first.
- This relay action can be two-way, as in the case of a long-distance phone call.

- Another use of the satellite is when, as is the case with television broadcasts, the ground station's uplink is then down linked over a wide region, so that it may be received by many different customers possessing compatible equipment.
- Still another use for satellites is observation, wherein the satellite is equipped with cameras or various sensors, and it merely downlinks any information it picks up from its vantage point.

The Ground Station:

- This is the earth segment.
- The ground station's job is two-fold. In the case of an uplink, or transmitting station, terrestrial data in the form of baseband signals, is passed through a baseband processor, an up converter, a high-powered amplifier, and through a parabolic dish antenna up to an orbiting satellite.
- In the case of a downlink, or receiving station, works in the reverse fashion as the uplink, ultimately
- Converting signals received through the parabolic antenna to base band signal.

<u>Cellular Telephony System:</u>

- A cellular telephone system provides a wireless connection to the terrestrial telephone network (PSTN: Public switch Telephone Network) for any user location within the radio range of the system.
- Cellular systems accommodate a large number of users over a large geographic area, within a limited frequency spectrum.
- Cellular radio system provides high quality service that is often comparable to that of the landline telephone systems.
- High capacity is achieved by limiting the coverage of each base station transmitter to a small geographic area called a cell so that the same radio channels may be reused by another base station located some distance away.
- A sophisticated switching technique called a handoff enables a call to proceed uninterrupted when the user moves from one cell to another.
- The basic structure of cellular system is as below:
- The basic cellular system consists mobile stations, base stations and a mobile switching center(MSC).
- The mobile switching center is sometimes called a mobile telephone switching office (MTSO), since it is responsible for connecting all mobiles to the PSTN via central office(CO).
- Each user communicates via radio from a cellular telephone set to the cell site base station.
- This base station connected via telephone lines or microwave link to the mobile switching center.
- The MSC connects the user to the called party if the called party is land based, the connection is via the central office (CO) is the terrestrial telephone network.

- If the called party is mobile, the connection is made to the cellular site that covers the area in which the third party is located, using an available radio channel in the cell associated with the called party.
- If more channels are needed, the existing cell sizes are decreased, and additional small cells are inserted, so that existing channels can be reused more efficiently.
- The critical consideration is to design the cells for acceptable levels of a Co channel interference.
- As the mobile user travels from one cell to another, the MSC automatically switches the user to an available channel in the new cell and the telephone continues un-interrupted.
- The cellular concept has following advantages:
    - large subscriber capacity
    - Efficient use of the radio spectrum
    - Service to hand held portables, as well as vehicles.
    - High Quality telephone and data service to the mobile user at relatively low cost.

****

Chapter 7

Impairments, Error handling and Compression Techniques

Introduction

When a signal transmit over a communication channel, it is subject to different types of impairments because of imperfect characteristics of the channel. Therefore, the received and the transmitted signals are not the same. Outcome of the impairments are manifest in two different ways in analog and digital signals. These impairments introduce random modifications in analog signals leading to distortion. On the other hand, in case of digital signals, the impairments lead to error in the bit values.

The impairment can be broadly categorizes into the following three types:

- Attenuation and attenuation distortion
- Delay distortion
- Noise

Attenuation

- Strength of a signal falls off with distance over any transmission medium.
- Hence, a received signal must have sufficient strength so that the electronic circuitry in the receiver can detect the signal.
- The signal must maintain a level sufficiently higher than noise to be receive without error.
- Attenuation is an increasing concern of frequency.
- Hence, amplifiers must use that amplify high frequencies more than lower frequencies.
- Attenuation is measure in decibels (dB). It measures the relative strengths of two signals or one signal at two different point.

  Attenuation (dB) = 10log10 (P2/P1)

  P1 is power at sending end and P2 is power at receiving end.

Delay Distortion

- Distortion means that the signal changes its shape.
- Distortion can occur in a composite signal made of different frequencies.
- Each signal component has its own propagation speed through a medium and, therefore, its own delay in arriving at the final destination. Difference in delay may create a difference in phase.
- The shape of the composite signal is therefore not the same.

Noise

- Noise is another cause of impairments.
- As signal transmits through a channel, undesired signal in the form of noise is mixed up with the signal, along with the distortion introduced by the transmission media.
- Noise can be categorized into the following four types:
  - Thermal Noise

- ➢ Intermodulation Noise
- ➢ Cross talk
- ➢ Impulse Noise

❖ **The thermal noise** is due to thermal agitation of electrons in a conductor. It is distributed across the entire spectrum and that is why it is also known as white noise (as the frequency encompass over a broad range of frequencies).

❖ When more than one signal share a single transmission medium, **intermodulation noise** is generated. For example, two signals f1 and f2 will generate signals of frequencies (f1 + f2) and (f1 - f2), which may interfere with the signals of the same frequencies sent by the transmitter. Intermodulation noise is introduced due to nonlinearity present in any part of the communication system.

❖ **Cross talk** is a result of bunching several conductors together in a single cable. Signal carrying wires generate electromagnetic radiation, which is induced on other conductors because of close proximity of the conductors. While using telephone, it is a common experience to hear conversation of other people in the background. This is known as cross talk.

❖ **Impulse noise** is irregular pulses or noise spikes of short duration generated by phenomena like lightning, spark due to loose contact in electric circuits, etc. Impulse noise is a primary source of bit-errors in digital data communication. This kind of noise introduces burst errors.

Error Detection and Correction Techniques

When data is transmit from one device to another device, the system does not guarantee whether the data received by the device is identical to the data transmitted by another device. An Error is a situation when the message received at the receiver end is not identical to the message transmitted

Errors classified into two categories:

- Single-Bit Error
- Burst Error

Single- Bit Error

The only one bit of a given data unit is changed from 1 to 0 or from 0 to 1.

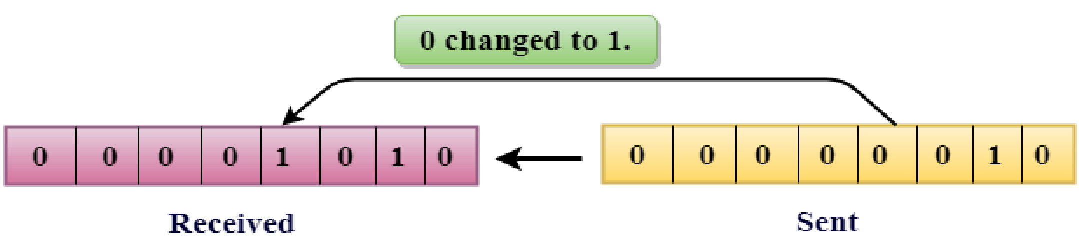


Figure 7.16: Single Bit Error

In the above figure, the message, which sent, was corrupt as single-bit, i.e., 0 bit is changed to one.

Single-Bit Error does not appear more likely in Serial Data Transmission. For example, Sender sends the data at 10 Mbps, this means that the bit lasts only for 1 ?s and for a single-bit error to occurred, a noise must be more than 1 ?s.

Single-Bit Error mainly occurs in Parallel Data Transmission. For example, if eight wires are used to send the eight bits of a byte, if one of the wire is noisy, then single-bit is corrupted per byte.

Burst Error

The two or more bits are changed from 0 to 1 or from 1 to 0 is known as Burst Error. The Burst Error is determined from the first corrupted bit to the last corrupted bit. The duration of noise in Burst Error is more than the duration of noise in Single-Bit. Burst Errors are most likely to occur in Serial Data Transmission. The number of affected bits depends on the duration of the noise and data rate.

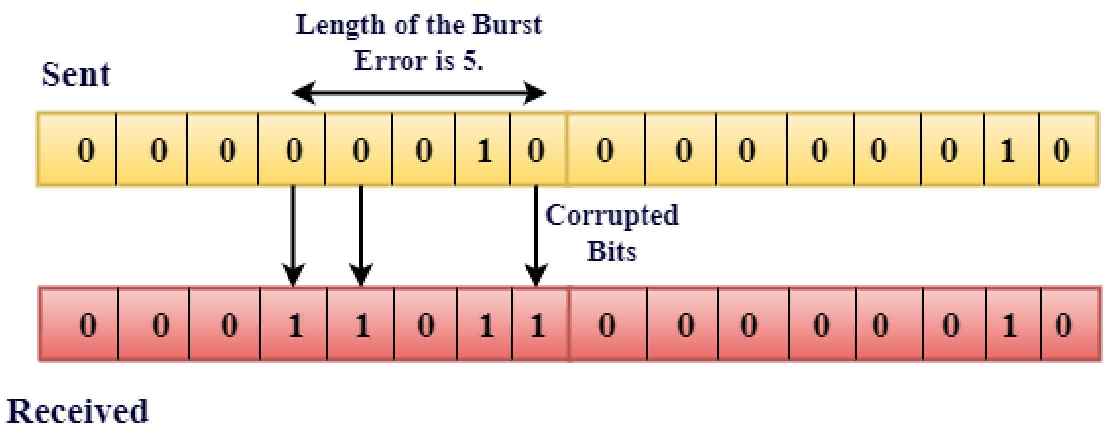


Figure 7.17: Burst Error

Error Detecting Techniques

The most popular Error Detecting Techniques are:

- Single parity check
- Two-dimensional parity check
- Checksum
- Cyclic redundancy check

❖ ***Single Parity Check***

- Single Parity checking is the simple mechanism and inexpensive to detect the errors.
- In this technique, a redundant bit known as a parity bit, which is append at the end of the data unit so that the number of 1's becomes even. Therefore, the total number of transmitted bits would be 9 bits.
- If the number of 1's bits is odd, then parity bit 1 is append and if the number of 1's bits is even, then parity bit 0 is appended at the end of the data unit.

- At the receiving end, the parity bit calculated from the received data bits and compared with the received parity bit.
- This technique generates the total number of 1's even, so it is known as even-parity checking.

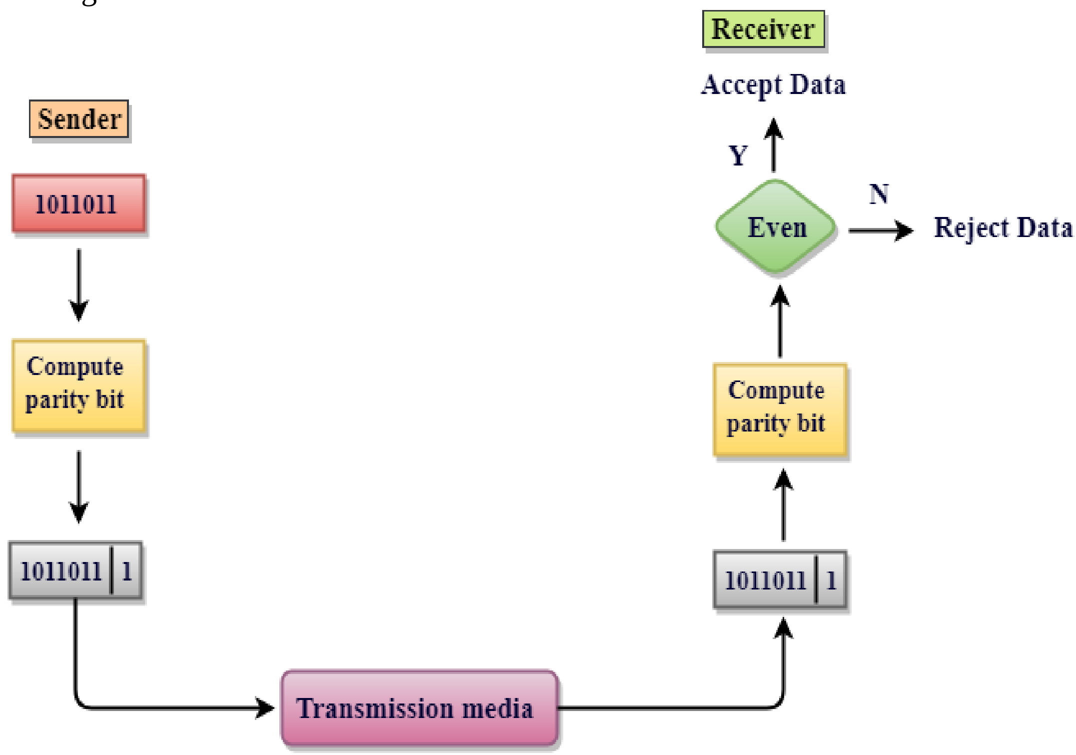


Figure 7.18: Even Parity Checking

Drawbacks of Single Parity Checking

- It can only detect single-bit errors, which are very rare.
- If two bits are interchange, then it cannot detect the errors.

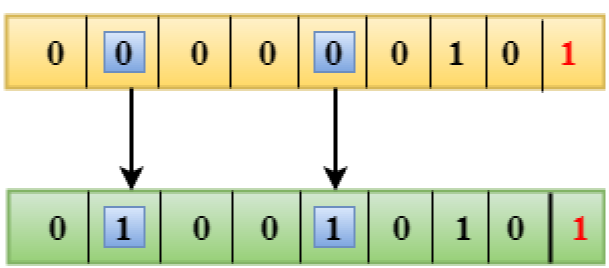


Figure 7.19: Drawback of Single Parity Checking

❖ **Two-Dimensional Parity Check**

- Performance can be improved by using Two-Dimensional Parity Check, which organizes the data in the form of a table.
- Parity check bits are compute for each row, which is equivalent to the single-parity check.

- In Two-Dimensional Parity check, a block of bits is divide into rows, and the redundant row of bits is added to the whole block.
- At the receiving end, the parity bits are compared with the parity bits computed from the received data.

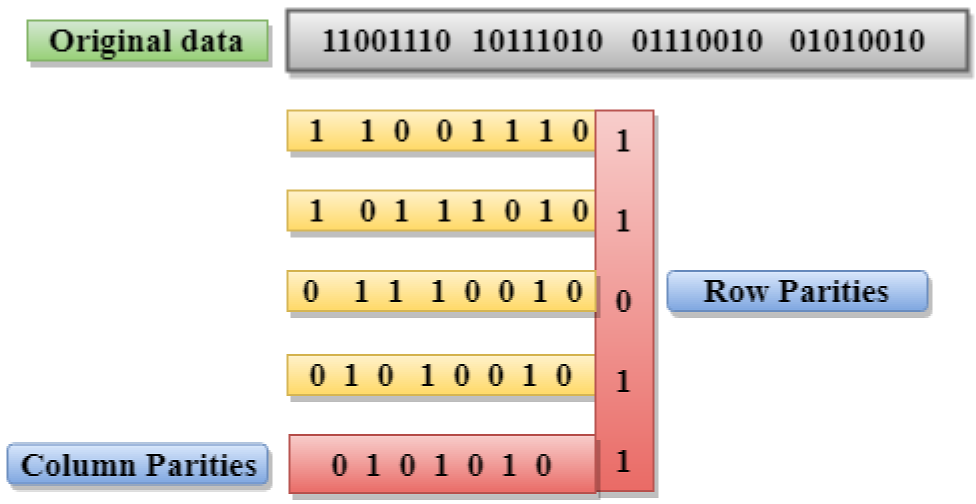


Figure 7.20: Two Dimensional Parity Check

Drawbacks Of 2D Parity Check

- If two bits in one data unit are corrupted and two bits the same position in another data unit are, also corrupted, then 2D Parity checker will not be able to detect the error.
- This technique cannot be used to detect the 4-bit errors or more in some cases.

❖ **Checksum**

- A Checksum is an error detection technique based on the concept of redundancy.
- It is divided into two parts:
  - ➢ **Checksum Generator**

A Checksum generated at the sending side. Checksum generator subdivides the data into equal segments of n bits each, and all these segments added together by using one's complement arithmetic. The sum is complemented and appended to the original data, known as checksum field. The extended data transmitted across the network.

Suppose L is the total sum of the data segments, then the checksum would be. L

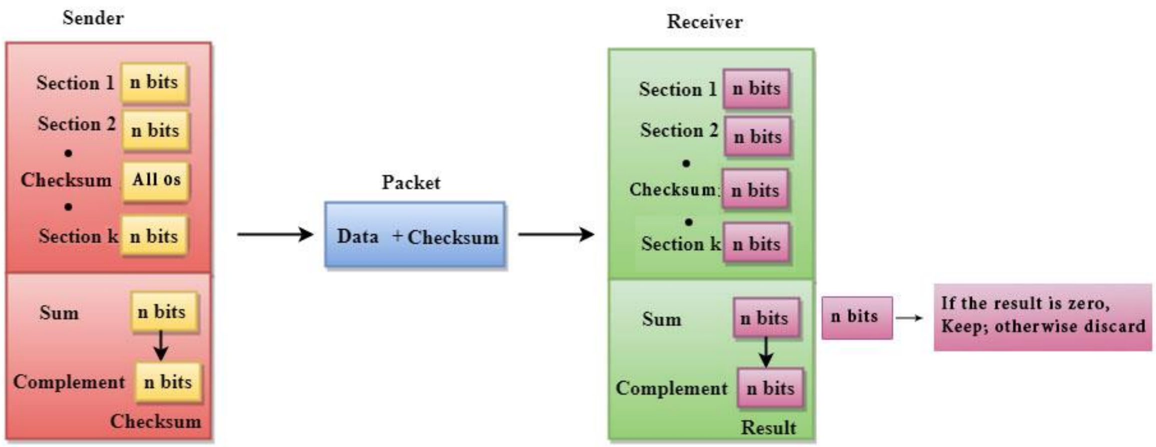


Figure 7.21: Check Sum Generator

The Sender follows the given steps:

- The block unit is divided into k sections, and each of n bits.
- All the k sections are added together by using one's complement to get the sum.
- The sum is complemented and it becomes the checksum field.
- The original data and checksum field are sent across the network.

➢ Checksum Checker

A Checksum is verified at the receiving side. The receiver subdivides the incoming data into equal segments of n bits each, and all these segments are added together, and then this sum is complemented. If the complement of the sum is zero, then the data is accepted otherwise data is rejected.

The Receiver follows the given steps:

- The block unit is divided into k sections and each of n bits.
- All the k sections are added together by using one's complement algorithm to get the sum.
- The sum is complemented.
- If the result of the sum is zero, then the data is accepted otherwise the data is discarded.

❖ Cyclic Redundancy Check (CRC)

CRC is a redundancy error technique used to determine the error.

Following are the steps used in CRC for error detection:

- In CRC technique, a string of n 0's is appended to the data unit, and this n number is less than the number of bits in a predetermined number, known as division which is n+1 bits.
- Secondly, the newly extended data is divide by a divisor using a process known as binary division. The remainder generated from this division known as CRC remainder.
- Thirdly, the CRC remainder replaces the appended 0's at the end of the original data. This newly generated unit is send to the receiver.
- The receiver receives the data followed by the CRC remainder. The receiver will treat this whole unit as a single unit, and it is divide by the same divisor that was used to find the CRC remainder.

If the resultant of this division is zero which means that it has no error, and the data is accepted.

If the resultant of this division is not zero, which means that, the data consists of an error. Therefore, the data is discarded.

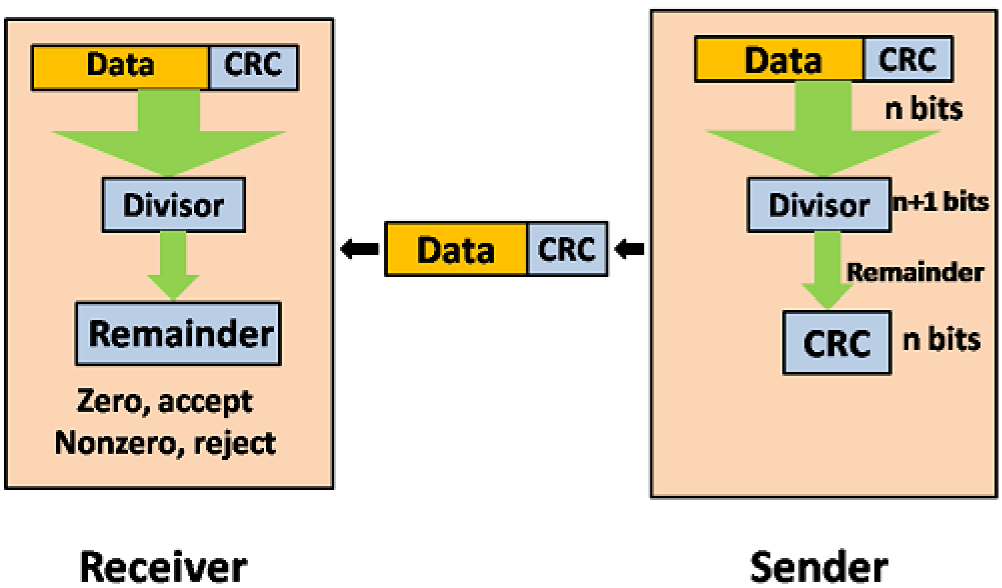


Figure 22.7: Cyclic Redundancy Check Overview

Let us understand this concept through an example:

Suppose the original data is 11100 and divisor is 1001.

CRC Generator

- A CRC generator uses a modulo-2 division. Firstly, three zeroes are appended at the end of the data as the length of the divisor is 4 and we know that the length of the string 0s to be appended is always one less than the length of the divisor.

- Now, the string becomes 11100000, and the resultant string is divided by the divisor 1001.
- The remainder generated from the binary division is known as CRC remainder. The generated value of the CRC remainder is 111.
- CRC remainder replaces the appended string of 0s at the end of the data unit, and the final string would be 11100111 which is sent across the network.

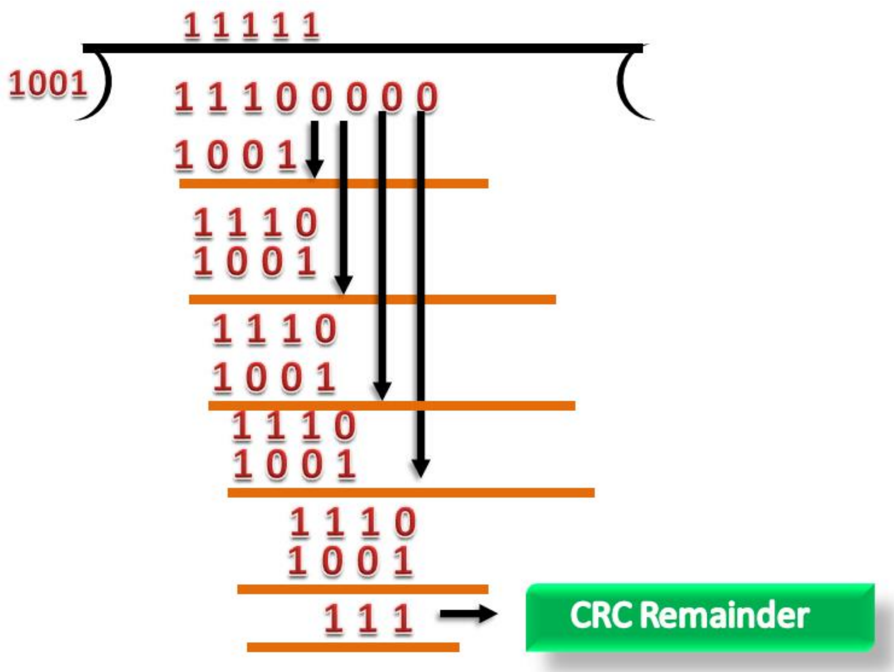


Figure 7.23: CRC Redundant Bit Generator

CRC Checker

- The functionality of the CRC checker is similar to the CRC generator.
- When the string 11100111 is received at the receiving end, then CRC checker performs the modulo-2 division.
- A string is divided by the same divisor, i.e., 1001.
- In this case, CRC checker generates the remainder of zero. Therefore, the data is accepted.

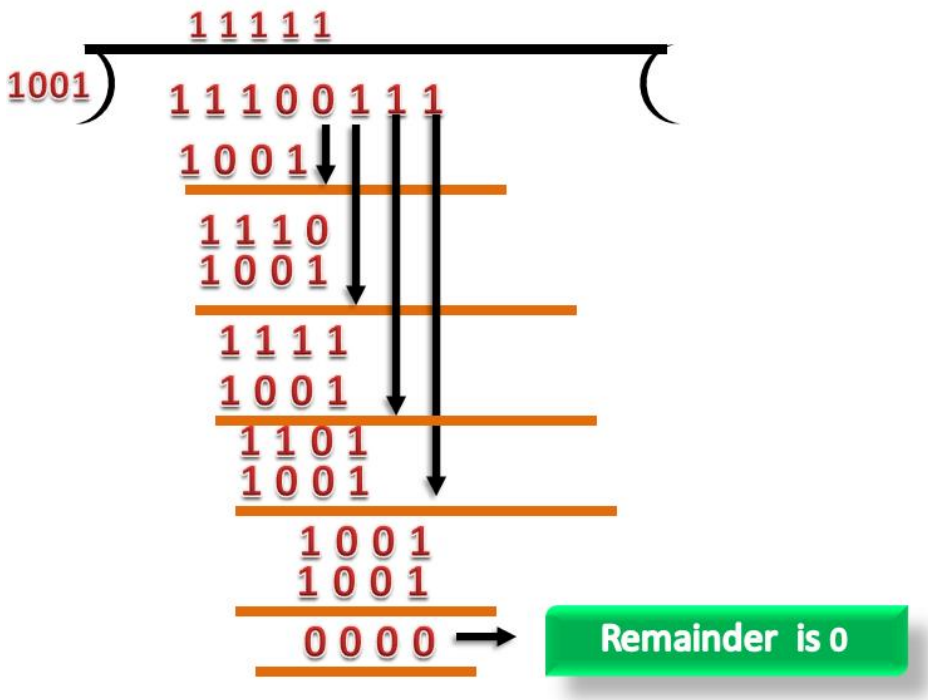


Figure 7.24: CRC Message Checker

Error Correction Technique

- Error Correction codes are used to detect and correct the errors when data is transmitted from the sender to the receiver.
- Error Correction can be handle in two ways:
  - **Backward error correction:** Once the error is discover, the receiver requests the sender to retransmit the entire data unit.
  - **Forward error correction:** In this case, the receiver uses the error-correcting code, which automatically corrects the errors.
- A single additional bit can detect the error, but cannot correct it.
- For correcting the errors, one has to know the exact position of the error.
- **For example,** If we want to calculate a single-bit error, the error correction code will determine which one of seven bits is in error. To achieve this, we have to add some additional redundant bits.

  Suppose r is the number of redundant bits and d is the total number of the data bits. The number of redundant bits r can be calculated by using the formula:

  2r>=d+r+1

  The value of r is calculated by using the above formula. For example, if the value of d is 4, then the possible smallest value that satisfies the above relation would be 3.

  To determine the position of the bit which is in error, a technique developed by R.W Hamming is Hamming code which can be applied to any length of the data unit and uses the relationship between data units and redundant units.

❖ Hamming Code

- Parity bits: The bit which is appended to the original data of binary bits so that the total number of 1s is even or odd.
- Even parity: To check for even parity, if the total number of 1s is even, then the value of the parity bit is 0. If the total number of 1s occurrences is odd, then the value of the parity bit is 1.
- Odd Parity: To check for odd parity, if the total number of 1s is even, then the value of parity bit is 1. If the total number of 1s is odd, then the value of parity bit is 0.

Algorithm of Hamming code:

- An information of 'd' bits are added to the redundant bits 'r' to form d + r.
- The location of each of the (d + r) digits is assigned a decimal value.
- The 'r' bits are placed in the positions 1,2,.....2^{k-1}.
- At the receiving end, the parity bits are recalculated. The decimal value of the parity bits determines the position of an error.

Let us understand the concept of Hamming code through an example:

Suppose the original data is 1010, which is to be sent.

- Total number of data bits 'd' = 4
- Number of redundant bits r : 2r >= d+r+1
- 2r>= 4+r+1
- Therefore, the value of r is 3 that satisfies the above relation.
- Total number of bits = d + r = 4+3 = 7;

Determining the position of the redundant bits

The number of redundant bits is 3. The three bits are represented by r1, r2, r4. The position of the redundant bits is calculated with corresponds to the raised power of 2. Therefore, their corresponding positions are 1, 21, 22.

The position of r1 = 1

The position of r2 = 2

The position of r4 = 4

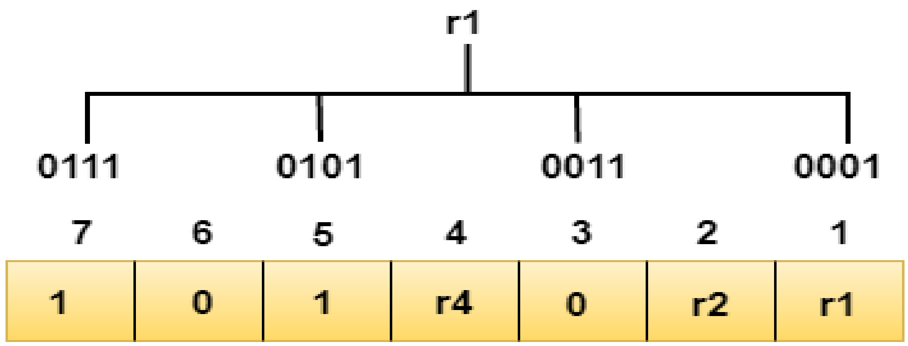
Representation of Data on the addition of parity bits:

7	6	5	4	3	2	1
1	0	1	r4	0	r2	r1

Determining the Parity bits

- Determining the **r1** bit

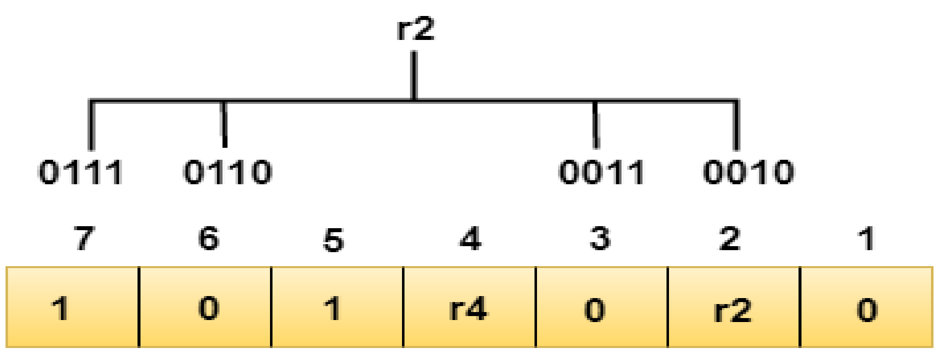
  The r1 bit is calculated by performing a parity check on the bit positions whose binary representation includes 1 in the first position.



  We observe from the above figure that the bit positions that includes 1 in the first position are 1, 3, 5, 7. Now, we perform the even-parity check at these bit positions. The total number of 1 at these bit positions corresponding to r1 is even, therefore, the value of the r1 bit is 0.

- Determining r2 bit

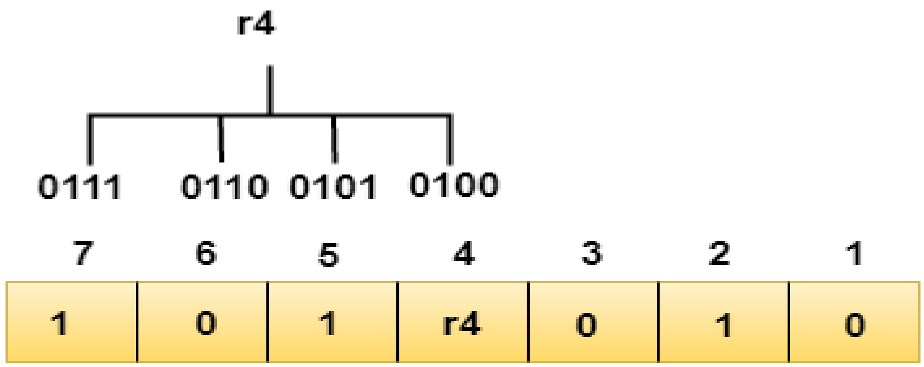
  The r2 bit is calculated by performing a parity check on the bit positions whose binary representation includes 1 in the second position.



  We observe from the above figure that the bit positions that includes 1 in the second position are 2, 3, 6, 7. Now, we perform the even-parity check at these bit positions. The total number of 1 at these bit positions corresponding to r2 is odd, therefore, the value of the r2 bit is 1.

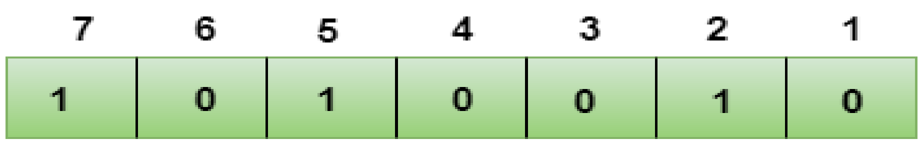
- Determining r4 bit

  The r4 bit is calculated by performing a parity check on the bit positions whose binary representation includes 1 in the third position.



We observe from the above figure that the bit positions that includes 1 in the third position are 4, 5, 6, 7. Now, we perform the even-parity check at these bit positions. The total number of 1 at these bit positions corresponding to r4 is even, therefore, the value of the r4 bit is 0.

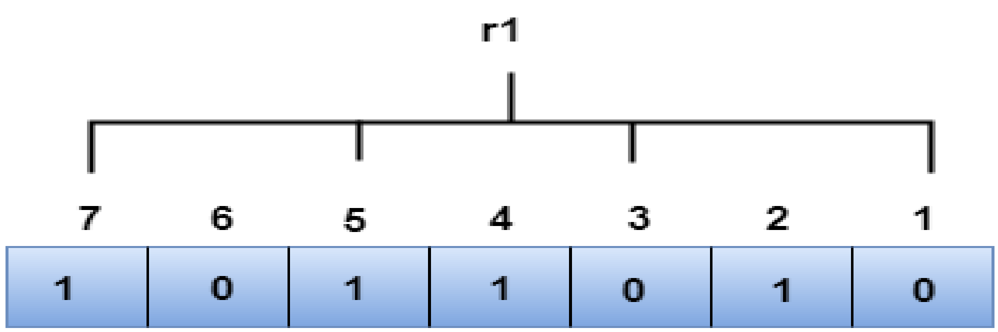
Data transferred is given below:



Suppose the 4th bit is changed from 0 to 1 at the receiving end, then parity bits are recalculated.

- R1 bit

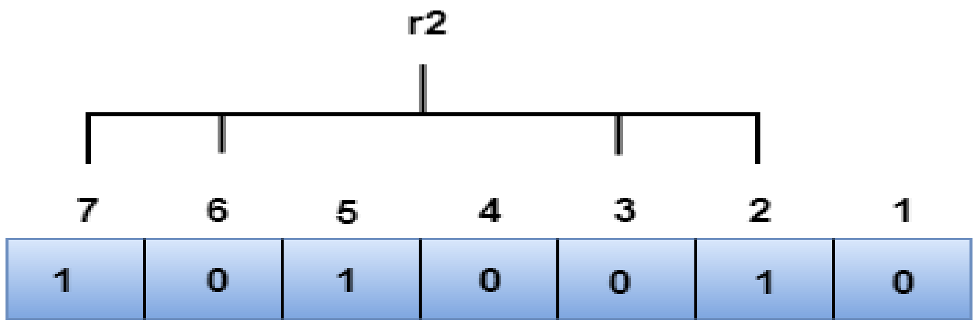
The bit positions of the r1 bit are 1,3,5,7



We observe from the above figure that the binary representation of r1 is 1100. Now, we perform the even-parity check, the total number of 1s appearing in the r1 bit is an even number. Therefore, the value of r1 is 0.

- R2 bit

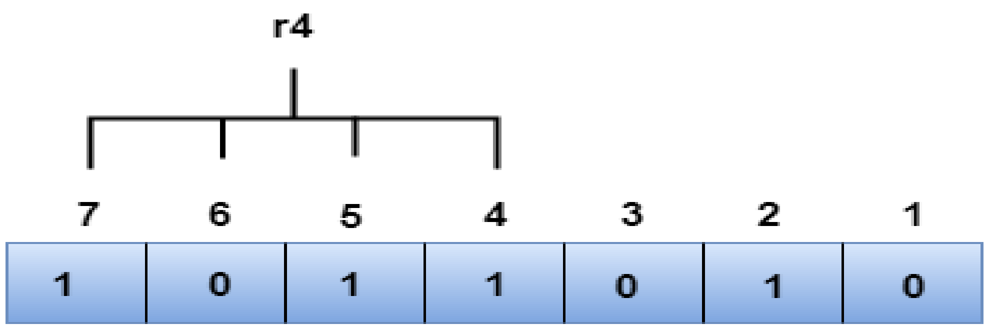
The bit positions of r2 bit are 2,3,6,7.



We observe from the above figure that the binary representation of r2 is 1001. Now, we perform the even-parity check, the total number of 1s appearing in the r2 bit is an even number. Therefore, the value of r2 is 0.

- R4 bit

The bit positions of r4 bit are 4,5,6,7.



We observe from the above figure that the binary representation of r4 is 1011. Now, we perform the even-parity check, the total number of 1s appearing in the r4 bit is an odd number. Therefore, the value of r4 is 1.

The binary representation of redundant bits, i.e., r4r2r1 is 100, and its corresponding decimal value is 4. Therefore, the error occurs in a 4^{th} bit position. The bit value must be changed from 1 to 0 to correct the error.

Data Compression

DC stands for Data Compression. DC is a digital signal process in which data to be transmitted is compressed to reduce the storage amount in bits. In other words, you can say that data storage space is reduced than usual after applying DC. Data transmission greatly reduces data storage space and transmission capacity. It is also known as source coding or bit-rate reduction. Database management system, backup utilities, etc. use data compression method widely. There are many file compression methods but ZIP and ARC are mostly known file formats.

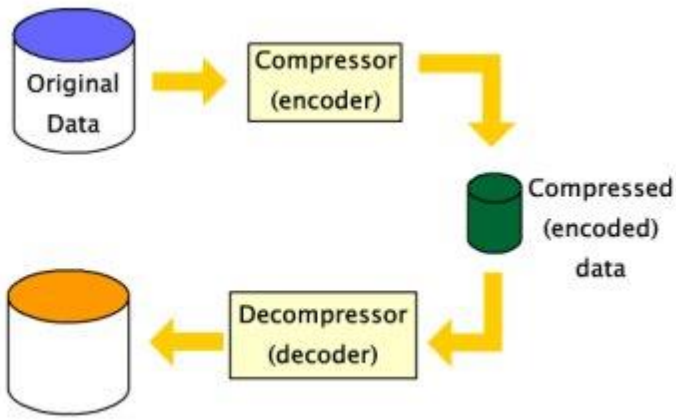


Figure 7.25: DC

Two Parts of DC

- Lossy: In lossy compression method, some part of data is deleted or lost. Because it identifies and then delete the unnecessary information before transmission.
- Lossless: In lossless compression method, compression is done through identifying and eliminating any statistical redundancy. For instance, when we encode a data source before transmitting it, its size is effectively reduced and data remains intact and unchanged.

Benefits of Data Compression

- Faster file transfer: It improves the speed of file transfer, as less bandwidth is required to download the compressed files.
- More Storage Capacity: It allows you store more files in the available storage space, e.g. Lossless compression can reduce a file to 50% of its original size.
- Reduces Cost: It allows you reduce the cost of storing data as after compression, you can store more files in the given storage space.
- Reduces latency: On tape, the small file images can be scanned faster to reach a specific file that reduces latency.

Difference between Lossy Compression and Lossless Compression:

1. Lossy compression is the method, which eliminate the data, which is not noticeable. While Lossless Compression does not eliminate the data which is not noticeable.
2. In Lossy compression, a file does not restore or rebuilt in its original form. While in Lossless Compression, A file can be restored in its original form.
3. In Lossy compression, Data's quality is compromised. However, Lossless Compression does not compromise the data's quality.
4. Lossy compression reduces the size of data. However, Lossless Compression does not reduce the size of data.
5. Algorithms used in Lossy compression are: Transform coding, Discrete Cosine Transform, Discrete Wavelet Transform, fractal compression etc. Algorithms

used in Lossless compression are Run Length Encoding, Lempel-Ziv-Welch, Huffman Coding, Arithmetic encoding etc.

6. Lossy compression is used in Images, audio, video. Lossless Compression used in Text, images, sound.
7. Lossy compression has more data-holding capacity. Lossless Compression has less data-holding capacity than Lossy compression technique.

Chapter 8

Data Link Control and Protocol

Data Control Link

Data Link Control is the service provided by the Data Link Layer to provide reliable data transfer over the physical medium. For example, In the half-duplex transmission mode, one device can only transmit the data at a time. If both the devices at the end of the links transmit the data simultaneously, they will collide and leads to the loss of the information. The Data link layer provides the coordination among the devices so that no collision occurs.

The Data link layer provides three functions:

- Line discipline
- Flow Control
- Error Control

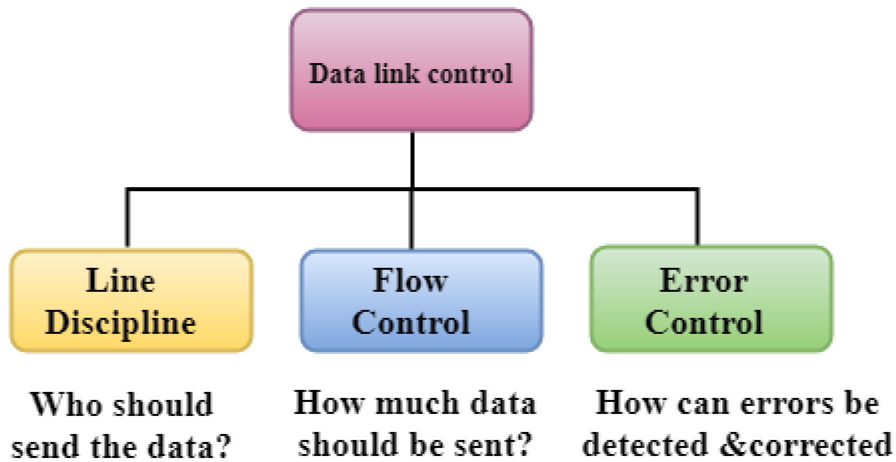


Figure 8.26: Data Link Control Functions

❖ **Line Discipline**

Line Discipline is a functionality of the Data link layer that provides the coordination among the link systems. It determines which device can send, and when it can send the data.

Line Discipline can be achieved in two ways:

- ENQ/ACK
- Poll/select

➢ **END/ACK**

END/ACK stands for Enquiry/Acknowledgement is use when there is no wrong receiver available on the link and having a dedicated path between the two devices so that the device capable of receiving the transmission is the intended one.

END/ACK coordinates which device will start the transmission and whether the recipient is ready or not.

Working of END/ACK

The transmitter transmits the frame called an Enquiry (ENQ) asking whether the receiver is available to receive the data or not.

The receiver responses either with the positive acknowledgement (ACK) or with the negative acknowledgement (NACK) where positive acknowledgement means that the receiver is ready to receive the transmission and negative acknowledgement means that the receiver is unable to accept the transmission.

Following are the responses of the receiver:

- If the response to the ENQ is positive, the sender will transmit its data, and once all of its data has been transmit, the device finishes its transmission with an EOT (END-of-Transmission) frame.
- If the response to the ENQ is negative, then the sender disconnects and restarts the transmission at another time.
- If the response is neither negative nor positive, the sender assumes that the ENQ frame was lost during the transmission and makes three attempts to establish a link before giving up.

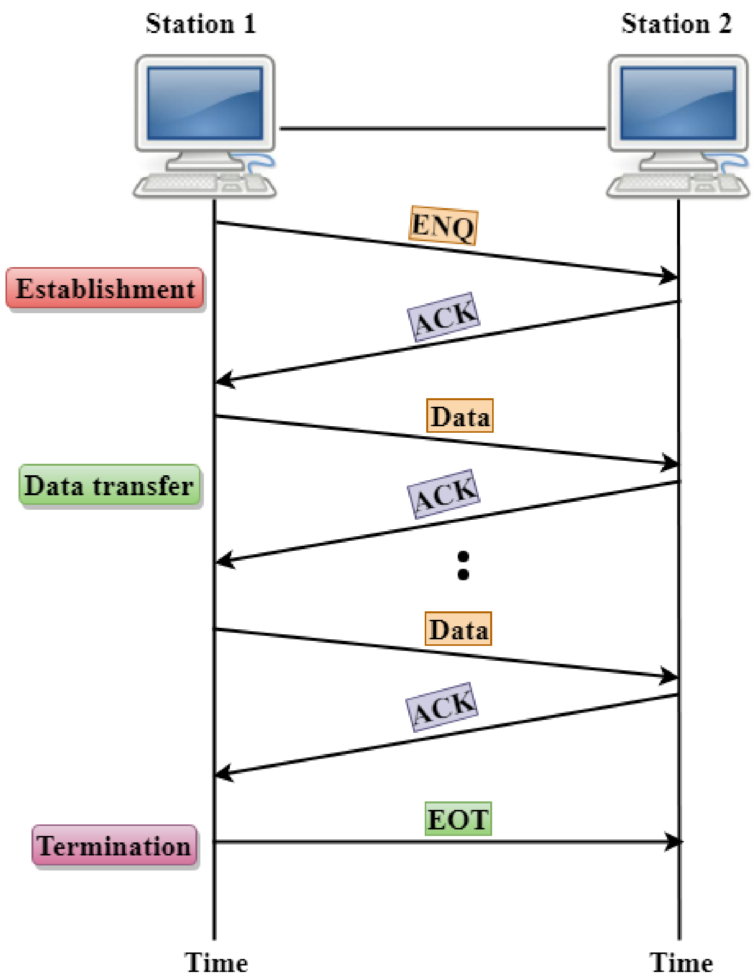


Figure 8.27: Working END/ACK

❖ Poll/Select

The Poll/Select method of line discipline works with those topologies where one device is designated as a primary station and other devices are secondary stations.

Working of Poll/Select

- In this, the primary device and multiple secondary devices consist of a single transmission line, and all the exchanges are made through the primary device even though the destination is a secondary device.
- The primary device has control over the communication link, and the secondary device follows the instructions of the primary device.
- The primary device determines which device is allowed to use the communication channel. Therefore, we can say that it is an initiator of the session.

- If the primary device wants to receive the data from the secondary device, it asks the secondary device that they anything to send, this process is known as polling.
- If the primary device wants to send some data to the secondary device, then it tells the target secondary to get ready to receive the data, this process is known as selecting.

Select

- The select mode is used when the primary device has something to send.
- When the primary device wants to send some data, then it alerts the secondary device for the upcoming transmission by transmitting a Select (SEL) frame, one field of the frame includes the address of the intended secondary device.
- When the secondary device receives the SEL frame, it sends an acknowledgement that indicates the secondary ready status.
- If the secondary device is ready to accept the data, then the primary device sends two or more data frames to the intended secondary device. Once the data has been transmitted, the secondary sends an acknowledgement specifies that the data has been received.

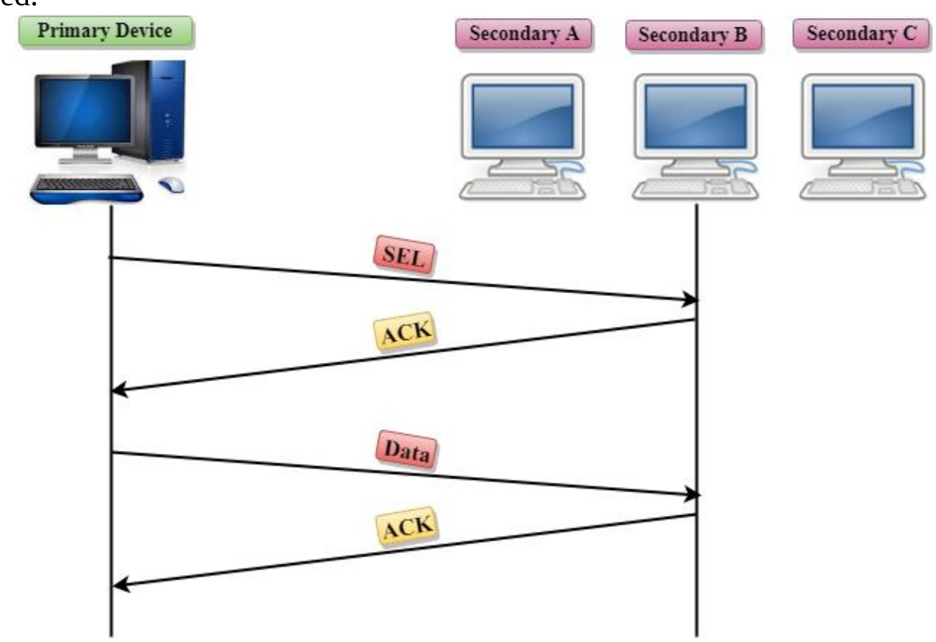


Figure 8.28: Select

Poll

- The Poll mode is used when the primary device wants to receive some data from the secondary device.
- When a primary device wants to receive the data, then it asks each device whether it has anything to send.
- Firstly, the primary asks (poll) the first secondary device, if it responds with the NACK (Negative Acknowledgement) means that it has nothing to send. Now, it approaches the second secondary device, it responds with the ACK means that it has the data to

send. The secondary device can send more than one frame one after another or sometimes it may be required to send ACK before sending each one, depending on the type of the protocol being used.

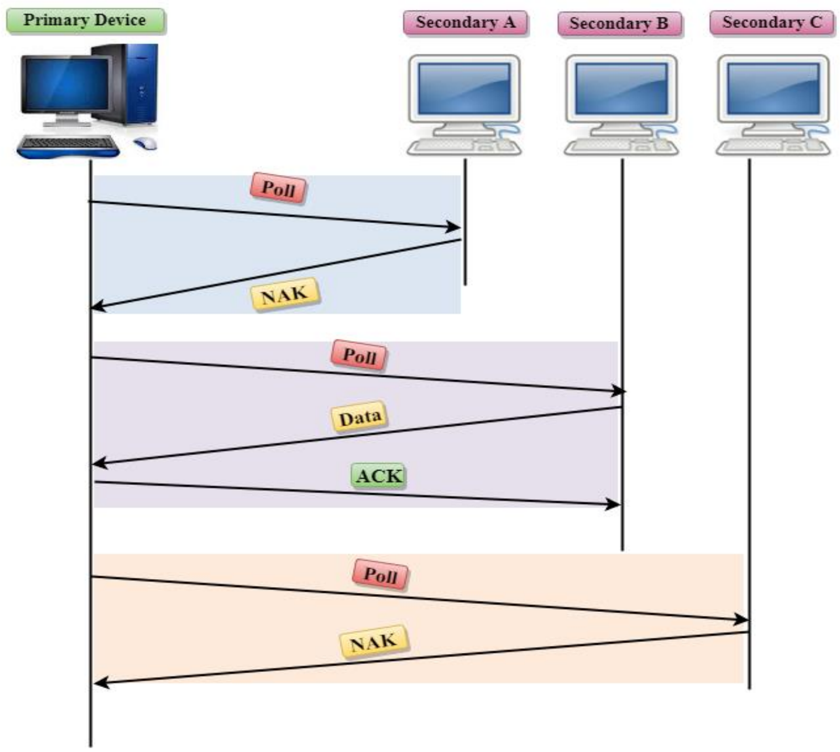


Figure 8.29: Poll

❖ <u>Flow Control</u>

- A set of procedures tells the sender how much data it can transmit before the data overwhelms the receiver.
- The receiving device has limited speed and limited memory to store the data. Therefore, the receiving device must be able to inform the sending device to stop the transmission temporarily before the limits are reached.
- It requires a buffer, a block of memory for storing the information until they are processed.

Two methods have been developed to control the flow of data:

- Stop-and-wait
- Sliding window

Stop-and-wait

- In the Stop-and-wait method, the sender waits for an acknowledgement after every frame it sends.

- When acknowledgement is received, then only next frame is sent. The process of alternately sending and waiting of a frame continues until the sender transmits the EOT (End of transmission) frame.

Advantage of Stop-and-wait

- The Stop-and-wait method is simple as each frame is checked and acknowledged before the next frame is sent.

Disadvantage of Stop-and-wait

- Stop-and-wait technique is inefficient to use as each frame must travel across all the way to the receiver, and an acknowledgement travels all the way before the next frame is sent. Each frame sent and received uses the entire time needed to traverse the link.

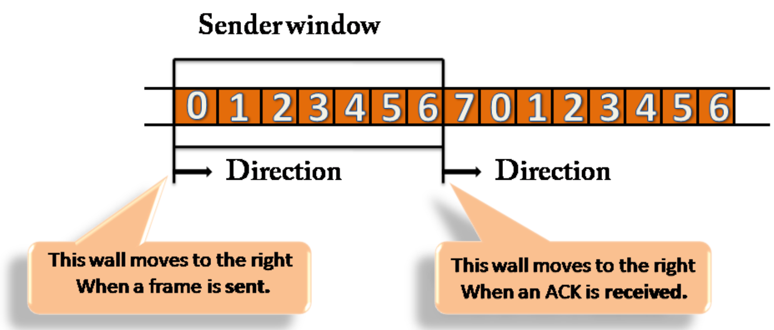
Sliding Window

- The Sliding Window is a method of flow control in which a sender can transmit the several frames before getting an acknowledgement.
- In Sliding Window Control, multiple frames can be sent one after another, due to which capacity of the communication channel can be utilized efficiently.
- A single ACK acknowledge multiple frames.
- Sliding Window refers to imaginary boxes at both the sender and receiver end.
- The window can hold the frames at either end, and it provides the upper limit on the number of frames that can be transmitted before the acknowledgement.
- Frames can be acknowledged even when the window is not completely filled.
- The window has a specific size in which they are numbered as modulo-n means that they are numbered from 0 to n-1. For example, if n = 8, the frames are numbered from 0,1,2,3,4,5,6,7,0,1,2,3,4,5,6,7,0,1........
- The size of the window is represented as n-1. Therefore, maximum n-1 frames can be sent before acknowledgement.
- When the receiver sends the ACK, it includes the number of the next frame that it wants to receive. For example, to acknowledge the string of frames ending with frame number 4, the receiver will send the ACK containing the number 5. When the sender sees the ACK with the number 5, it got to know that the frames from 0 through 4 have been received.

Sender Window

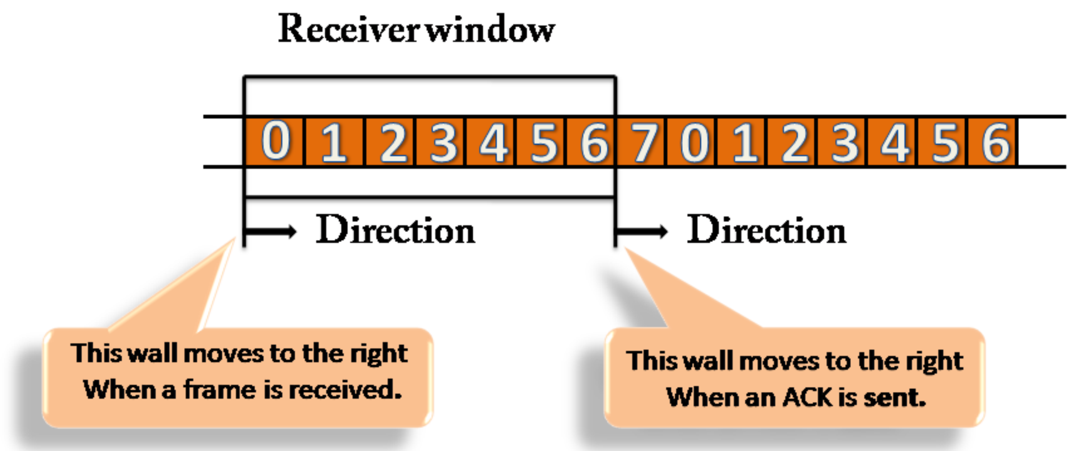
- At the beginning of a transmission, the sender window contains n-1 frames, and when they are sent out, the left boundary moves inward shrinking the size of the window. For example, if the size of the window is w if three frames are sent out, then the number of frames left out in the sender window is w-3.
- Once the ACK has arrived, then the sender window expands to the number which will be equal to the number of frames acknowledged by ACK.

- For example, the size of the window is 7, and if frames 0 through 4 have been sent out and no acknowledgement has arrived, then the sender window contains only two frames, i.e., 5 and 6. Now, if ACK has arrived with a number 4 which means that 0 through 3 frames have arrived undamaged and the sender window is expanded to include the next four frames. Therefore, the sender window contains six frames (5,6,7,0,1,2).



Receiver Window

- At the beginning of transmission, the receiver window does not contain n frames, but it contains n-1 spaces for frames.
- When the new frame arrives, the size of the window shrinks.
- The receiver window does not represent the number of frames received, but it represents the number of frames that can be received before an ACK is sent. For example, the size of the window is w, if three frames are received then the number of spaces available in the window is (w-3).
- Once the acknowledgement is sent, the receiver window expands by the number equal to the number of frames acknowledged.
- Suppose the size of the window is 7 means that the receiver window contains seven spaces for seven frames. If the one frame is received, then the receiver window shrinks and moving the boundary from 0 to 1. In this way, window shrinks one by one, so window now contains the six spaces. If frames from 0 through 4 have sent, then the window contains two spaces before an acknowledgement is sent.



❖ **<u>Error Control</u>**

Error Control is a technique of error detection and retransmission.

Categories of Error Control:

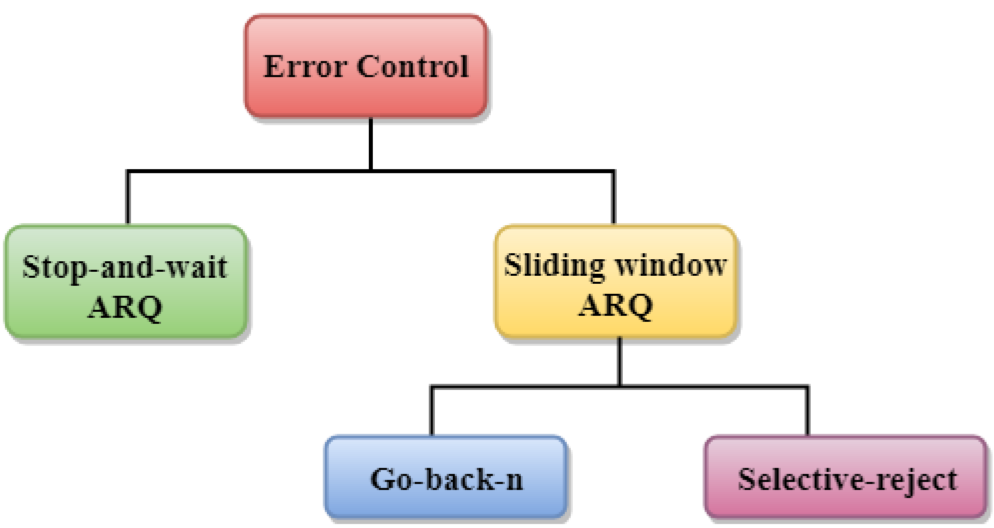


Figure 8.30: Categories of Error Control

Stop-and-wait ARQ

Stop-and-wait ARQ is a technique used to retransmit the data in case of damaged or lost frames. This technique works on the principle that the sender will not transmit the next frame until it receives the acknowledgement of the last transmitted frame.

Four features are required for the retransmission:

- The sending device keeps a copy of the last transmitted frame until the acknowledgement is received. Keeping the copy allows the sender to retransmit the data if the frame is not received correctly.
- Both the data frames and the ACK frames are numbered alternately 0 and 1 so that they can be identified individually. Suppose data 1 frame acknowledges the data 0 frame means that the data 0 frame has been arrived correctly and expects to receive data 1 frame.
- If an error occurs in the last transmitted frame, then the receiver sends the NAK frame which is not numbered. On receiving the NAK frame, sender retransmits the data.
- It works with the timer. If the acknowledgement is not received within the allotted time, then the sender assumes that the frame is lost during the transmission, so it will retransmit the frame.

Two possibilities of the retransmission:

- Damaged Frame: When the receiver receives a damaged frame, i.e., the frame contains an error, then it returns the NAK frame. For example, when the data 0 frame is sent, and then the receiver sends the ACK 1 frame means that the data 0 has arrived correctly, and transmits the data 1 frame. The sender transmits the next frame: data 1. It reaches undamaged, and the receiver returns ACK 0. The sender transmits the next frame: data 0. The receiver reports an error and returns the NAK frame. The sender retransmits the data 0 frame.
- Lost Frame: Sender is equipped with the timer and starts when the frame is transmitted. Sometimes the frame has not arrived at the receiving end so that it can be acknowledged neither positively nor negatively. The sender waits for acknowledgement until the timer goes off. If the timer goes off, it retransmits the last transmitted frame.

Sliding Window ARQ

Sliding Window ARQ is a technique used for continuous transmission error control.

Three Features used for retransmission:

- In this case, the sender keeps the copies of all the transmitted frames until they have been acknowledged. Suppose the frames from 0 through 4 have been transmitted, and the last acknowledgement was for frame 2, the sender has to keep the copies of frames 3 and 4 until they receive correctly.
- The receiver can send either NAK or ACK depending on the conditions. The NAK frame tells the sender that the data have been received damaged. Since the sliding window is a continuous transmission mechanism, both ACK and NAK must be numbered for the identification of a frame. The ACK frame consists of a number that represents the next frame which the receiver expects to receive. The NAK frame consists of a number that represents the damaged frame.
- The sliding window ARQ is equipped with the timer to handle the lost acknowledgements. Suppose then n-1 frames have been sent before receiving any acknowledgement. The sender waits for the acknowledgement, so it starts the timer and

waits before sending any more. If the allotted time runs out, the sender retransmits one or all the frames depending upon the protocol used.

Two protocols used in sliding window ARQ:

- Go-Back-n ARQ: In Go-Back-N ARQ protocol, if one frame is lost or damaged, then it retransmits all the frames after which it does not receive the positive ACK.

Three possibilities can occur for retransmission:

- Damaged Frame: When the frame is damaged, then the receiver sends a NAK frame.

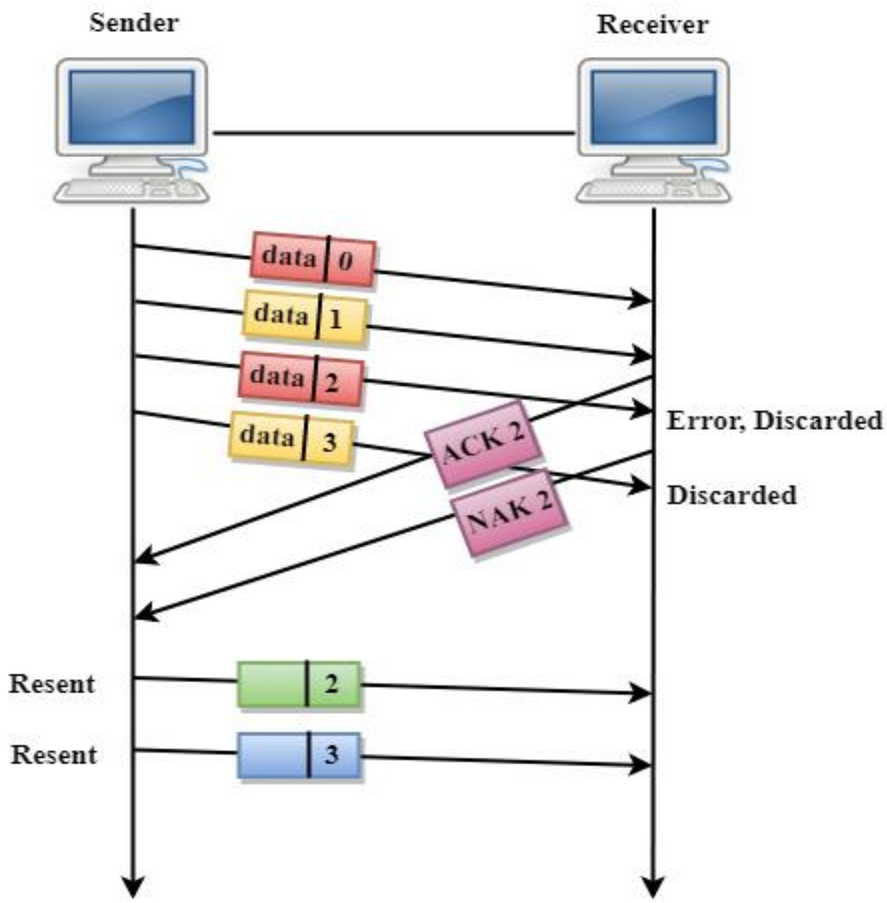


Figure8.31: Sliding window ARQ

In the above figure, three frames have been transmitted before an error discovered in the third frame. In this case, ACK 2 has been returned telling that the frames 0,1 have been received successfully without any error. The receiver discovers the error in data 2 frame, so it returns the NAK 2 frame. The frame 3 is also discarded as it is transmitted after the damaged frame. Therefore, the sender retransmits the frames 2,3.

- **Lost Data Frame:** In Sliding window protocols, data frames are sent sequentially. If any of the frames is lost, then the next frame arrive at the receiver is out of sequence. The receiver checks the sequence number of each of the frame, discovers the frame that has been skipped, and returns the NAK for the missing frame. The sending device retransmits the frame indicated by NAK as well as the frames transmitted after the lost frame.

- **Lost Acknowledgement:** The sender can send as many frames as the windows allow before waiting for any acknowledgement. Once the limit of the window is reached, the sender has no more frames to send; it must wait for the acknowledgement. If the acknowledgement is lost, then the sender could wait forever. To avoid such situation, the sender is equipped with the timer that starts counting whenever the window capacity is reached. If the acknowledgement has not been received within the time limit, then the sender retransmits the frame since the last ACK.

Selective-Reject ARQ

- Selective-Reject ARQ technique is more efficient than Go-Back-n ARQ.
- In this technique, only those frames are retransmitted for which negative acknowledgement (NAK) has been received.
- The receiver storage buffer keeps all the damaged frames on hold until the frame in error is correctly received.
- The receiver must have an appropriate logic for reinserting the frames in a correct order.
- The sender must consist of a searching mechanism that selects only the requested frame for retransmission.

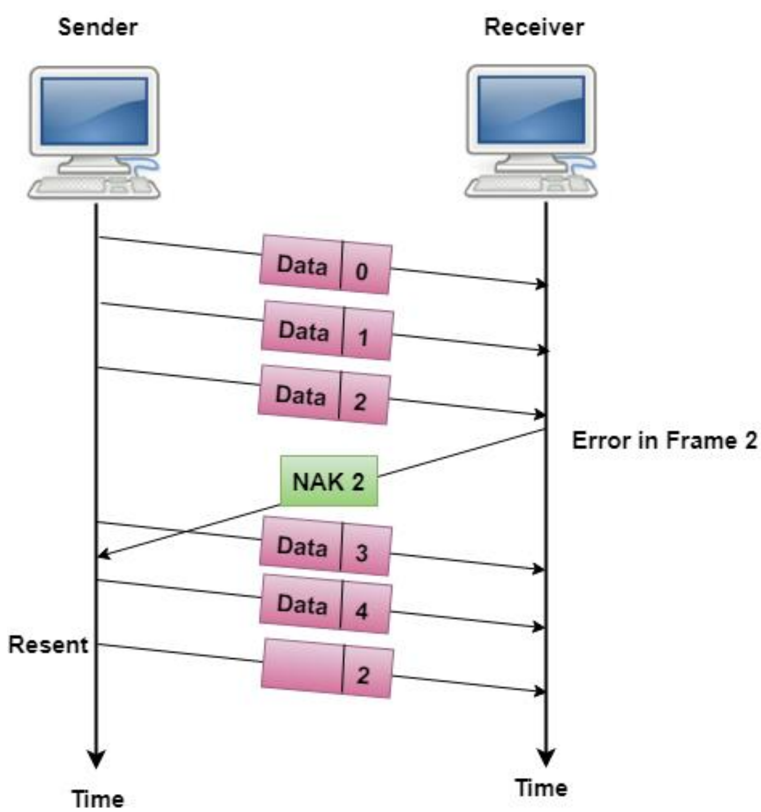


Figure 8.32: Go n Back

Chapter 9

Multiplexing and Switching

Multiplexing

- Multiplexing is a technique used to combine and send the multiple data streams over a single medium. The process of combining the data streams is known as multiplexing and hardware used for multiplexing is known as a multiplexer.
- Multiplexing is achieved by using a device called Multiplexer (MUX) that combines n input lines to generate a single output line. Multiplexing follows many-to-one, i.e., n input lines and one output line.
- Demultiplexing is achieved by using a device called DE multiplexer (DEMUX) available at the receiving end. DEMUX separates a signal into its component signals (one input and n outputs). Therefore, we can say that demultiplexing follows the one-to-many approach.

Why Multiplexing?

- The transmission medium is used to send the signal from sender to receiver. The medium can only have one signal at a time.
- If there are multiple signals to share one medium, then the medium must be divided in such a way that each signal is given some portion of the available bandwidth. For example: If there are 10 signals and bandwidth of medium is100 units, then the 10 unit is shared by each signal.
- When multiple signals share the common medium, there is a possibility of collision. Multiplexing concept is used to avoid such collision.
- Transmission services are very expensive.

History of Multiplexing

- Multiplexing technique is widely used in telecommunications in which several telephone calls are carried through a single wire.
- Multiplexing originated in telegraphy in the early 1870s and is now widely used in communication.
- **George Owen Squier** developed the telephone carrier multiplexing in 1910.

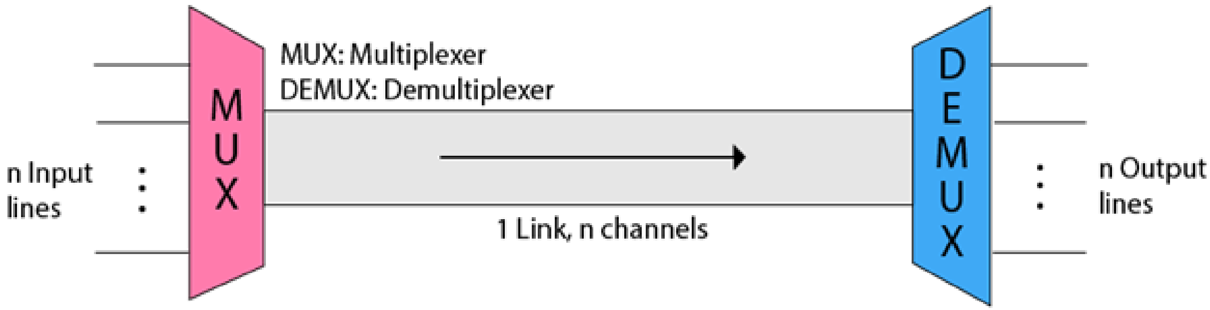


Figure 9.33: Concept of Multiplexing

- The 'n' input lines are transmitted through a multiplexer and multiplexer combines the signals to form a composite signal.
- The composite signal is passed through a Demultiplexer and demultiplexer separates a signal to component signals and transfers them to their respective destinations.

Advantages of Multiplexing:

- More than one signal can be sent over a single medium.
- The bandwidth of a medium can be utilized effectively.

Multiplexing Techniques

Multiplexing techniques can be classified as:

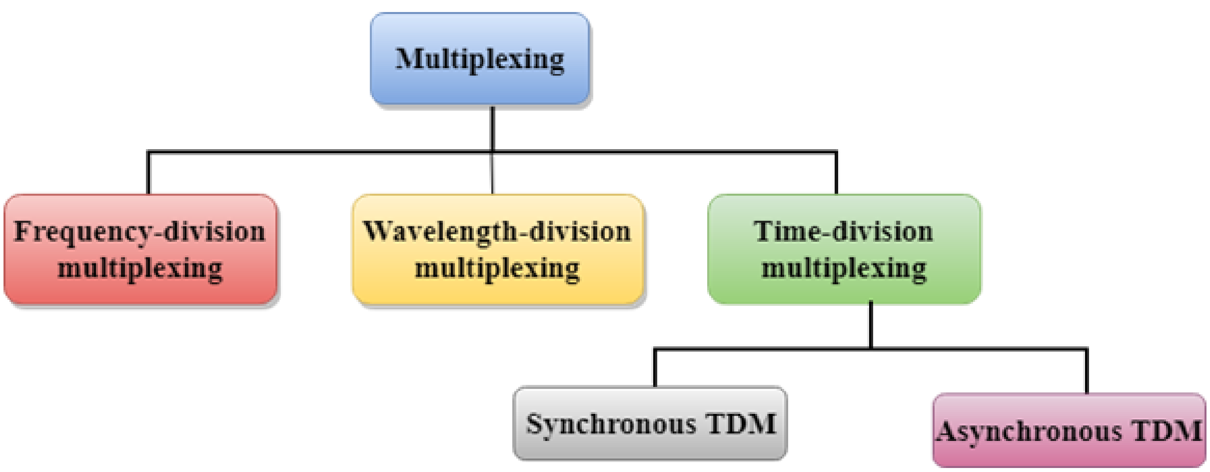
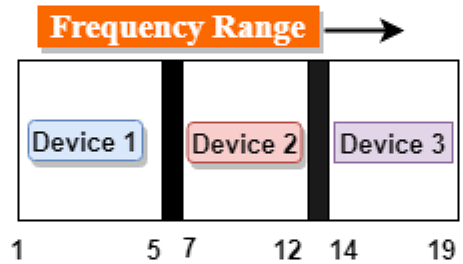


Figure 9.34: Multiplexing Techniques

i. Frequency-division Multiplexing (FDM)

- It is an analog technique.
- Frequency Division Multiplexing is a technique in which the available bandwidth of a single transmission medium is subdivided into several channels.



- In the above diagram, a single transmission medium is subdivided into several frequency channels, and each frequency channel is given to different devices. Device 1 has a frequency channel of range from 1 to 5.
- The input signals are translated into frequency bands by using modulation techniques, and they are combined by a multiplexer to form a composite signal.
- The main aim of the FDM is to subdivide the available bandwidth into different frequency channels and allocate them to different devices.
- Using the modulation technique, the input signals are transmitted into frequency bands and then combined to form a composite signal.
- The carriers which are used for modulating the signals are known as sub-carriers. They are represented as f1,f2..fn.
- FDM is mainly used in radio broadcasts and TV networks.

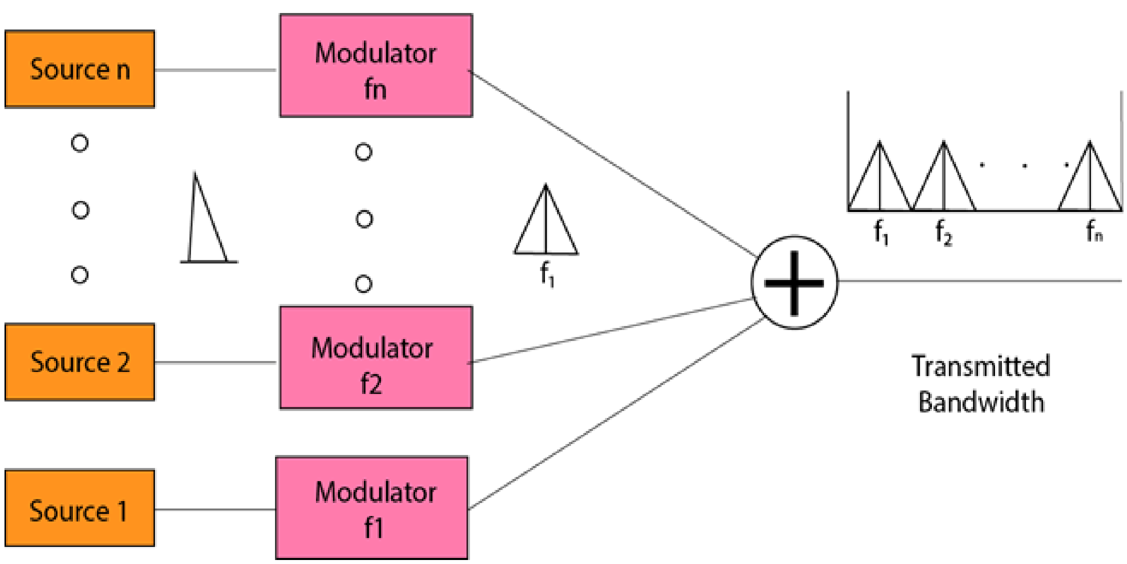


Figure 9.35: Frequency Division Multiplexing

- Advantages Of FDM:
    - FDM is used for analog signals.
    - FDM process is very simple and easy modulation.
    - A Large number of signals can be sent through an FDM simultaneously.
    - It does not require any synchronization between sender and receiver.
- Disadvantages Of FDM:

- FDM technique is used only when low-speed channels are required.
- It suffers the problem of crosstalk.
- A Large number of modulators are required.
- It requires a high bandwidth channel.

- Applications Of FDM:
  - FDM is commonly used in TV networks.
  - It is used in FM and AM broadcasting. Each FM radio station has different frequencies, and they are multiplexed to form a composite signal. The multiplexed signal is transmitted in the air.

ii. Wavelength Division Multiplexing (WDM)

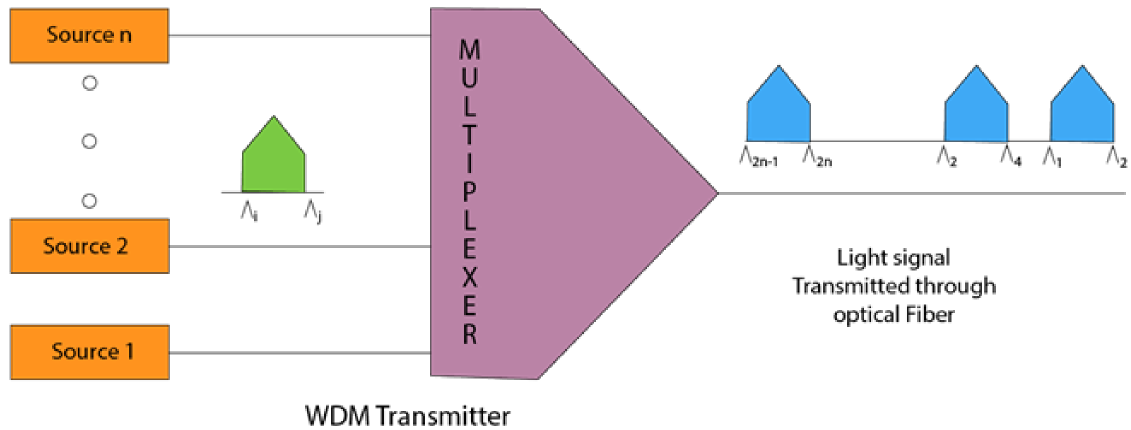
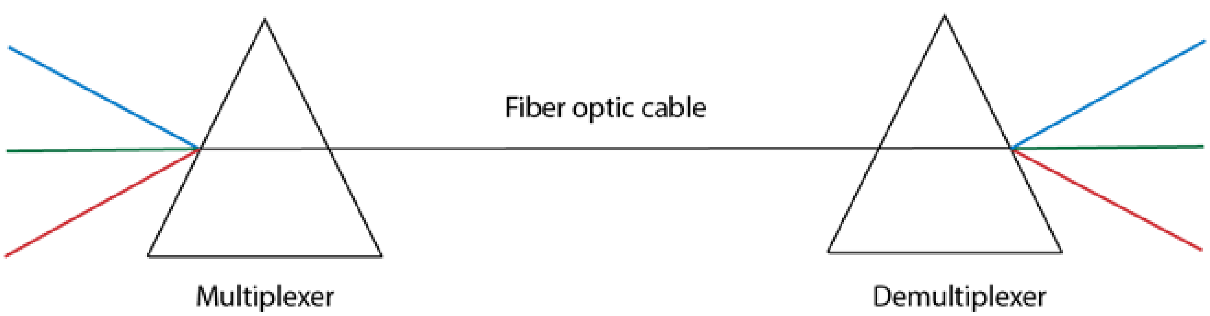


Figure 9.36: Wavelength Division Multiplexing

- Wavelength Division Multiplexing is same as FDM except that the optical signals are transmitted through the fiber optic cable.
- WDM is used on fiber optics to increase the capacity of a single fiber.
- It is used to utilize the high data rate capability of fiber optic cable.
- It is an analog multiplexing technique.
- Optical signals from different source are combined to form a wider band of light with the help of multiplexer.
- At the receiving end, demultiplexer separates the signals to transmit them to their respective destinations.
- Multiplexing and Demultiplexing can be achieved by using a prism.
- Prism can perform a role of multiplexer by combining the various optical signals to form a composite signal, and the composite signal is transmitted through a fiber optical cable.
- Prism also performs a reverse operation, i.e., demultiplexing the signal.



iii. Time Division Multiplexing

- It is a digital technique.
- In Frequency Division Multiplexing Technique, all signals operate at the same time with different frequency, but in case of Time Division Multiplexing technique, all signals operate at the same frequency with different time.

- In Time Division Multiplexing technique, the total time available in the channel is distributed among different users. Therefore, each user is allocated with different time interval known as a Time slot at which data is to be transmitted by the sender.
- A user takes control of the channel for a fixed amount of time.
- In Time Division Multiplexing technique, data is not transmitted simultaneously rather the data is transmitted one-by-one.
- In TDM, the signal is transmitted in the form of frames. Frames contain a cycle of time slots in which each frame contains one or more slots dedicated to each user.
- It can be used to multiplex both digital and analog signals but mainly used to multiplex digital signals.
- There are two types of TDM:
  - Synchronous TDM
  - Asynchronous TDM

Synchronous TDM

- A Synchronous TDM is a technique in which time slot is preassigned to every device.
- In Synchronous TDM, each device is given some time slot irrespective of the fact that the device contains the data or not.
- If the device does not have any data, then the slot will remain empty.
- In Synchronous TDM, signals are sent in the form of frames. Time slots are organized in the form of frames. If a device does not have data for a particular time slot, then the empty slot will be transmitted.
- The most popular Synchronous TDM are T-1 multiplexing, ISDN multiplexing, and SONET multiplexing.
- If there are n devices, then there are n slots.

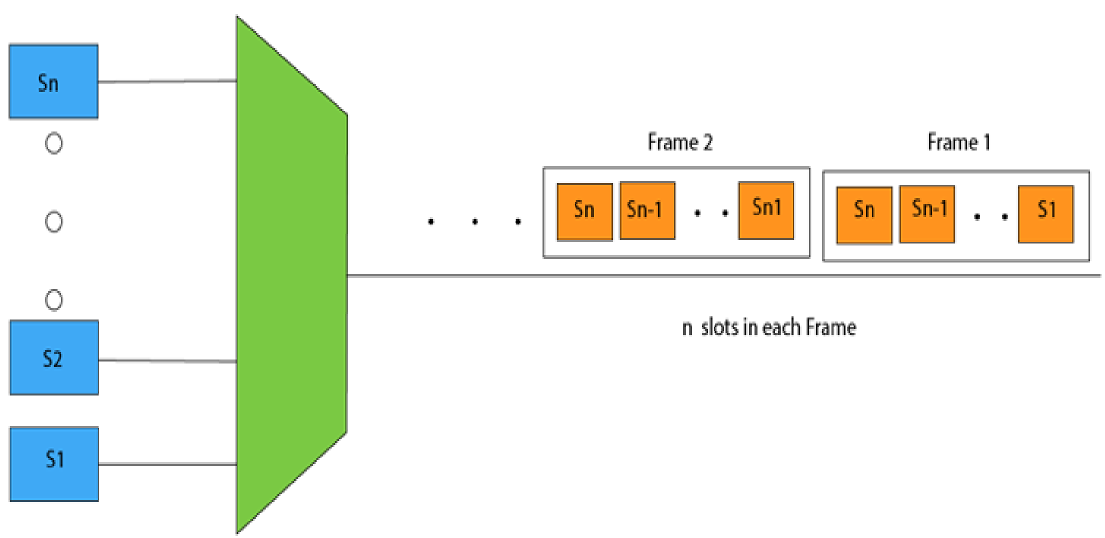


Figure 9.37: Concept of Synchronous TDM

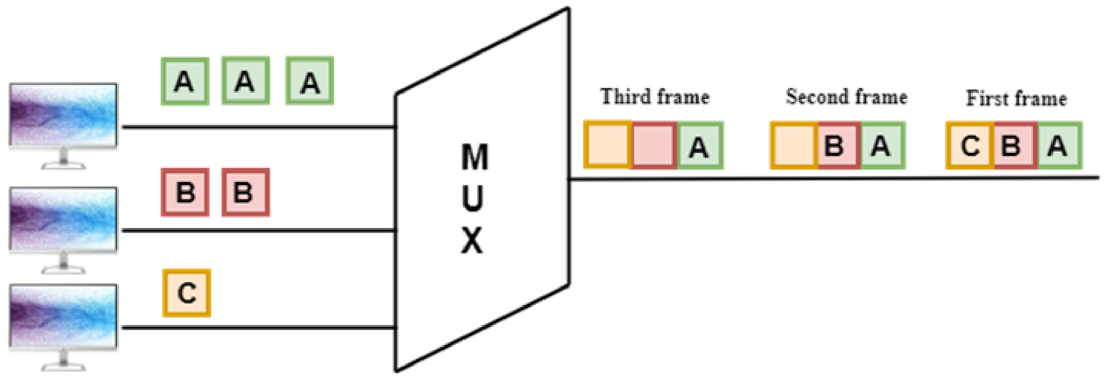


Figure 9.38: Synchronous TDM

In the above figure, the Synchronous TDM technique is implemented. Each device is allocated with some time slot. The time slots are transmitted irrespective of whether the sender has data to send or not.

Disadvantages of Synchronous TDM:

- The capacity of the channel is not fully utilized as the empty slots are also transmitted which is having no data. In the above figure, the first frame is completely filled, but in the last two frames, some slots are empty. Therefore, we can say that the capacity of the channel is not utilized efficiently.
- The speed of the transmission medium should be greater than the total speed of the input lines. An alternative approach to the Synchronous TDM is Asynchronous Time Division Multiplexing.

Asynchronous TDM

- An asynchronous TDM is also known as Statistical TDM.
- An asynchronous TDM is a technique in which time slots are not fixed as in the case of Synchronous TDM. Time slots are allocated to only those devices which have the data to send. Therefore, we can say that Asynchronous Time Division multiplexor transmits only the data from active workstations.
- An asynchronous TDM technique dynamically allocates the time slots to the devices.
- In Asynchronous TDM, total speed of the input lines can be greater than the capacity of the channel.
- Asynchronous Time Division multiplexor accepts the incoming data streams and creates a frame that contains only data with no empty slots.
- In Asynchronous TDM, each slot contains an address part that identifies the source of the data.

ADDRESS	DATA

- The difference between Asynchronous TDM and Synchronous TDM is that many slots in Synchronous TDM are unutilized, but in Asynchronous TDM, slots are fully utilized. This leads to the smaller transmission time and efficient utilization of the capacity of the channel.
- In Synchronous TDM, if there are n sending devices, then there are n time slots. In Asynchronous TDM, if there are n sending devices, then there are m time slots where m is less than n (m<n).
- The number of slots in a frame depends on the statistical analysis of the number of input lines.

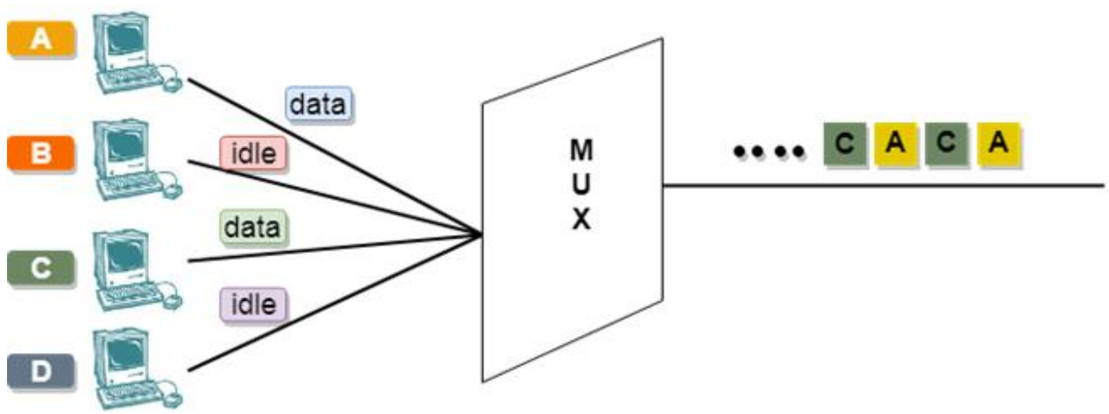
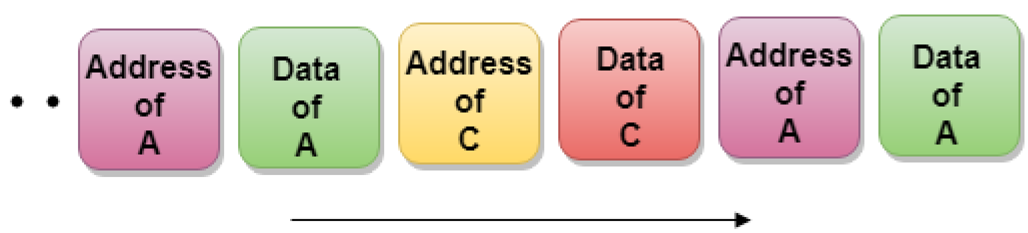


Figure 9.39: Concept of Asynchronous TDM

In the above diagram, there are 4 devices, but only two devices are sending the data, i.e., A and C. Therefore, the data of A and C are only transmitted through the transmission line.

Frame of above diagram can be represented as:



Transmission Line

The above figure shows that the data part contains the address to determine the source of the data.

Switching

- When a user accesses the internet or another computer network outside their immediate location, messages are sent through the network of transmission media. This technique of transferring the information from one computer network to another network is known as switching.
- Switching in a computer network is achieved by using switches. A switch is a small hardware device which is used to join multiple computers together with one local area network (LAN).
- Network switches operate at layer 2 (Data link layer) in the OSI model.
- Switching is transparent to the user and does not require any configuration in the home network.
- Switches are used to forward the packets based on MAC addresses.
- A Switch is used to transfer the data only to the device that has been addressed. It verifies the destination address to route the packet appropriately.
- It is operated in full duplex mode.
- Packet collision is minimum as it directly communicates between source and destination.
- It does not broadcast the message as it works with limited bandwidth.

Why is Switching Concept required?

Switching concept is developed because of the following reasons:

- **Bandwidth**: It is defined as the maximum transfer rate of a cable. It is a very critical and expensive resource. Therefore, switching techniques are used for the effective utilization of the bandwidth of a network.
- **Collision**: Collision is the effect that occurs when more than one device transmits the message over the same physical media, and they collide with each other. To overcome this problem, switching technology is implemented so that packets do not collide with each other.

Advantages of Switching:

- Switch increases the bandwidth of the network.
- It reduces the workload on individual PCs as it sends the information to only that device which has been addressed.
- It increases the overall performance of the network by reducing the traffic on the network.
- There will be less frame collision as switch creates the collision domain for each connection.

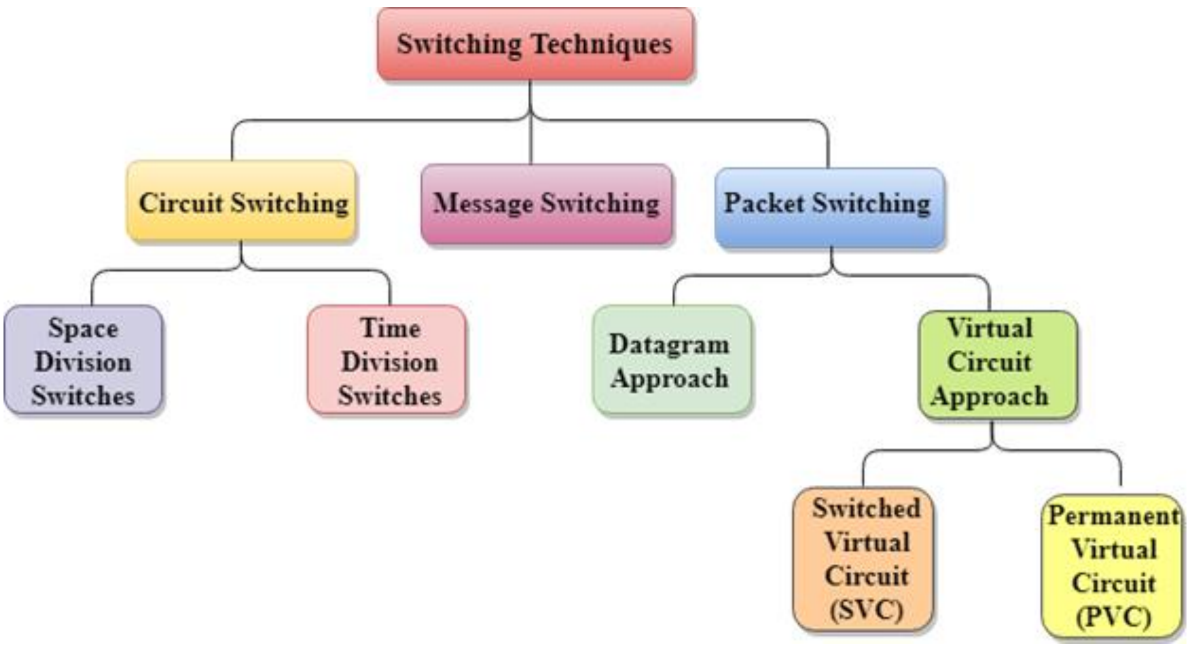
Disadvantages of Switching:

- A Switch is more expensive than network bridges.
- A Switch cannot determine the network connectivity issues easily.
- Proper designing and configuration of the switch are required to handle multicast packets.

Switching techniques

In large networks, there can be multiple paths from sender to receiver. The switching technique will decide the best route for data transmission. Switching technique is used to connect the systems for making one-to-one communication.

Classification of Switching Techniques



i. Circuit Switching

- Circuit switching is a switching technique that establishes a dedicated path between sender and receiver.
- In the Circuit Switching Technique, once the connection is established then the dedicated path will remain to exist until the connection is terminated.
- Circuit switching in a network operates in a similar way as the telephone works.
- A complete end-to-end path must exist before the communication takes place.
- In case of circuit switching technique, when any user wants to send the data, voice, video, a request signal is sent to the receiver then the receiver sends back the acknowledgment to ensure the availability of the dedicated path. After receiving the acknowledgment, dedicated path transfers the data.
- Circuit switching is used in public telephone network. It is used for voice transmission.
- Fixed data can be transferred at a time in circuit switching technology.

- Communication through circuit switching has 3 phases:
    - Circuit establishment
    - Data transfer
    - Circuit Disconnect

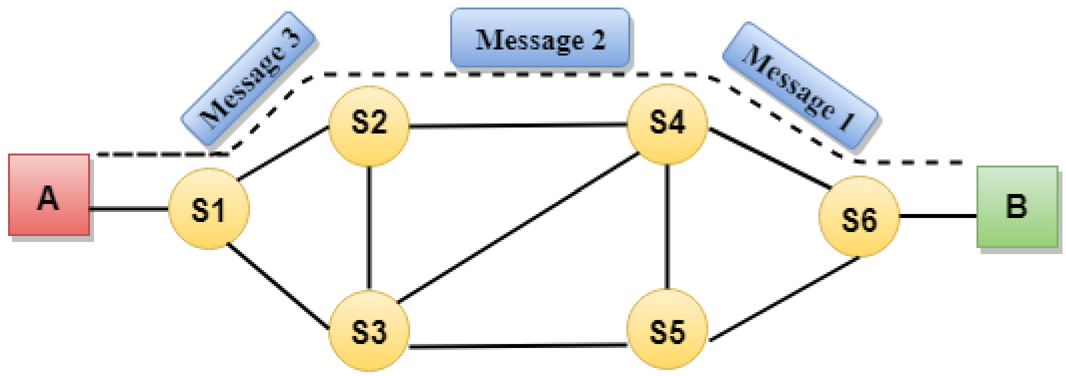


Figure 9.40: Circuit switching between two nodes

Circuit Switching can use either of the two technologies:

- Space Division Switches:
    - Space Division Switching is a circuit switching technology in which a single transmission path is accomplished in a switch by using a physically separate set of cross points.
    - Space Division Switching can be achieved by using crossbar switch. A crossbar switch is a metallic cross point or semiconductor gate that can be enabled or disabled by a control unit.
    - The Crossbar switch is made by using the semiconductor. For example, Xilinx crossbar switch using FPGAs.
    - Space Division Switching has high speed, high capacity, and non-blocking switches.
- Space Division Switches can be categorized in two ways:
    - Crossbar Switch
    - Multistage Switch

a. Crossbar Switch

The Crossbar switch is a switch that has n input lines and n output lines. The crossbar switch has n2 intersection points known as cross points.

Disadvantage of Crossbar switch:

The number of cross points increases as the number of stations is increased. Therefore, it becomes very expensive for a large switch. The solution to this is to use a multistage switch.

b. Multistage Switch

- Multistage Switch is made by splitting the crossbar switch into the smaller units and then interconnecting them.
- It reduces the number of cross points.
- If one path fails, then there will be an availability of another path.
- *Advantages Of Circuit Switching*:
    - In the case of Circuit Switching technique, the communication channel is dedicated.
    - It has fixed bandwidth.
- *Disadvantages Of Circuit Switching:*
    - Once the dedicated path is established, the only delay occurs in the speed of data transmission.
    - It takes a long time to establish a connection approx. 10 seconds during which no data can be transmitted.
    - It is more expensive than other switching techniques as a dedicated path is required for each connection.
    - It is inefficient to use because once the path is established and no data is transferred, then the capacity of the path is wasted.
    - In this case, the connection is dedicated therefore no other data can be transferred even if the channel is free.

ii. Message Switching

- Message Switching is a switching technique in which a message is transferred as a complete unit and routed through intermediate nodes at which it is stored and forwarded.
- In Message Switching technique, there is no establishment of a dedicated path between the sender and receiver.
- The destination address is appended to the message. Message Switching provides a dynamic routing as the message is routed through the intermediate nodes based on the information available in the message.
- Message switches are programmed in such a way so that they can provide the most efficient routes.
- Each and every node stores the entire message and then forward it to the next node. This type of network is known as store and forward network.
- Message switching treats each message as an independent entity.

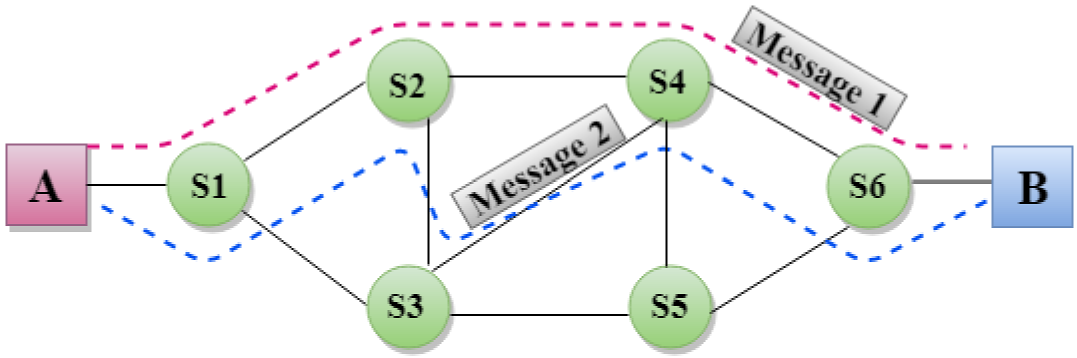


Figure 9.41: Concepts of Message Switching

- *Advantages Of Message Switching*
    - Data channels are shared among the communicating devices that improve the efficiency of using available bandwidth.
    - Traffic congestion can be reduced because the message is temporarily stored in the nodes.
    - Message priority can be used to manage the network.
    - The size of the message which is sent over the network can be varied. Therefore, it supports the data of unlimited size.
- *Disadvantages Of Message Switching*
    - The message switches must be equipped with sufficient storage to enable them to store the messages until the message is forwarded.
    - The Long delay can occur due to the storing and forwarding facility provided by the message switching technique.

iii. Packet Switching

- The packet switching is a switching technique in which the message is sent in one go, but it is divided into smaller pieces, and they are sent individually.
- The message splits into smaller pieces known as packets and packets are given a unique number to identify their order at the receiving end.
- Every packet contains some information in its headers such as source address, destination address and sequence number.
- Packets will travel across the network, taking the shortest path as possible.
- All the packets are reassembled at the receiving end in correct order.
- If any packet is missing or corrupted, then the message will be sent to resend the message.
- If the correct order of the packets is reached, then the acknowledgment message will be sent.

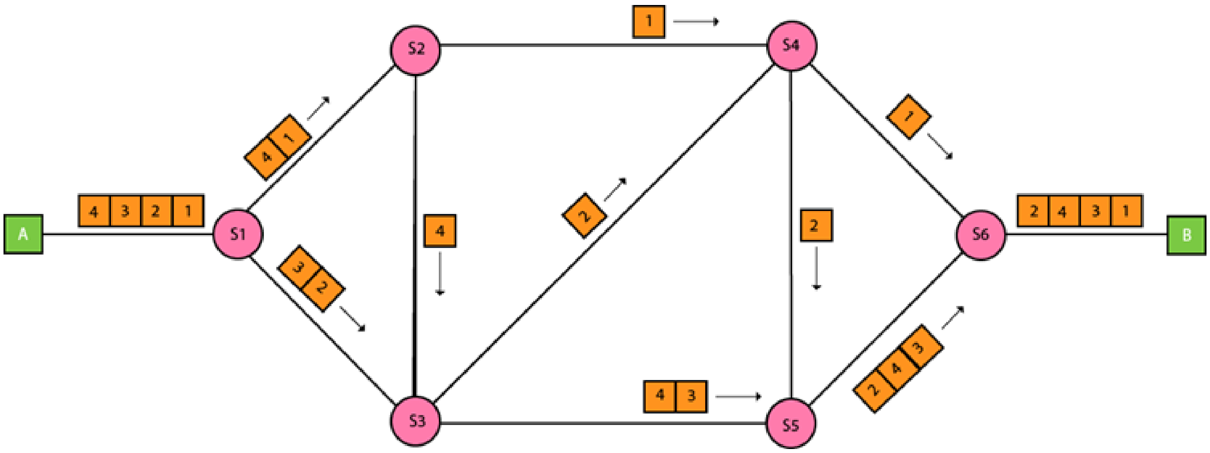


Figure 9.42: Concept of Packet Switching

- There are two approaches to Packet Switching:

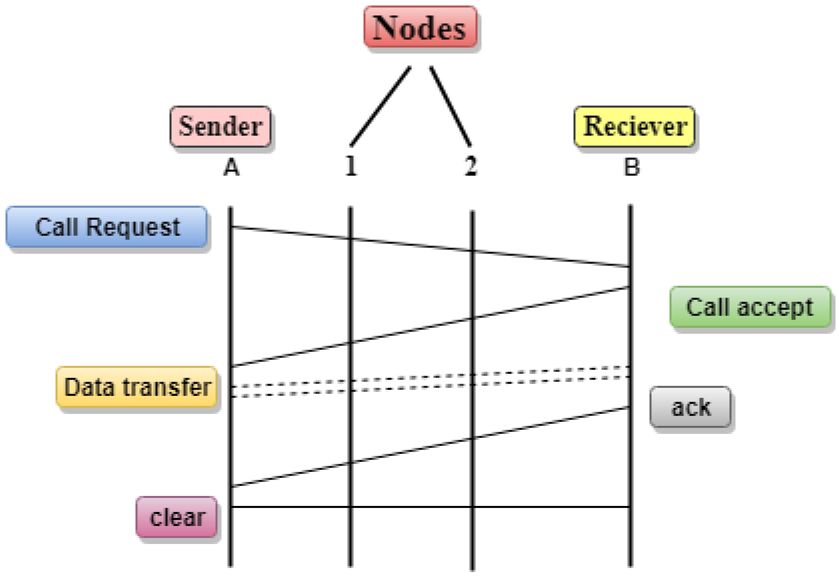
a. Datagram Packet switching:

- It is a packet switching technology in which packet is known as a datagram, is considered as an independent entity. Each packet contains the information about the destination and switch uses this information to forward the packet to the correct destination.
- The packets are reassembled at the receiving end in correct order.
- In Datagram Packet Switching technique, the path is not fixed.
- Intermediate nodes take the routing decisions to forward the packets.
- Datagram Packet Switching is also known as connectionless switching.

b. Virtual Circuit Switching

- Virtual Circuit Switching is also known as connection-oriented switching.
- In the case of Virtual circuit switching, a preplanned route is established before the messages are sent.
- Call request and call accept packets are used to establish the connection between sender and receiver.
- In this case, the path is fixed for the duration of a logical connection.

Let's understand the concept of virtual circuit switching through a diagram:



- In the above diagram, A and B are the sender and receiver respectively. 1 and 2 are the nodes.
- Call request and call accept packets are used to establish a connection between the sender and receiver.
- When a route is established, data will be transfer.
- After transmission of data, an acknowledgment signal sent by the receiver that the message has been received.
- If the user wants to terminate the connection, a clear signal is sent for the termination.

- *Advantages Of Packet Switching:*
  - Cost-effective: In packet switching technique, switching devices do not require massive secondary storage to store the packets, so cost is minimized to some extent. Therefore, we can say that the packet switching technique is a cost-effective technique.
  - Reliable: If any node is busy, then the packets can be rerouted. This ensures that the Packet Switching technique provides reliable communication.
  - Efficient: Packet Switching is an efficient technique. It does not require any established path prior to the transmission, and many users can use the same communication channel simultaneously, hence makes use of available bandwidth very efficiently.
- *Disadvantages Of Packet Switching:*
  - Packet Switching technique cannot be implemented in those applications that require low delay and high-quality services.
  - The protocols used in a packet switching technique are very complex and requires high implementation cost.

- If the network is overloaded or corrupted, then it requires retransmission of lost packets, it can also lead to the loss of critical information if errors are nor recovered.

Cellular Architecture

Cellular architecture is constituted of the following −

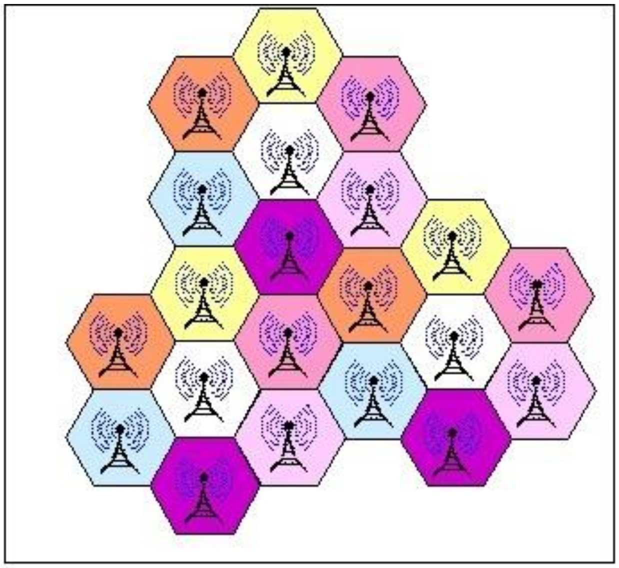
- A network of cells each with a base station.
- A packet switched network for communication between the base stations and mobile switching centers.
- The public switched telephone network to connect subscribers to the wider telephony network

Cellular Configuration

In all cellular systems, land area is divided into a number of cells each with its radio service. In AMPS the area is large which in digital services, the area is much smaller. Conventionally cells are hexagonal in shape. Each cell uses a frequency range that is not use by its adjacent cells. However, frequencies may be reused in non-adjacent cells.

At the center of each cell is a base station through which mobile phones and other mobile devices transmit data and voice.

The following diagram represents the configuration.



Mobile Switching Centre's

A mobile switching Centre (MSC) is a network switching subsystem of a cellular phone system. It is also called mobile telephone switching office (MTSO). All base stations are connected to an MSC.

The functions of MSC are −

- Call set-up and release.
- Routing of calls and messages sent via SMS.
- Managing conference calls and calls on hold.
- Fax services.
- Billing
- Interfacing with other networks like public switched telephone network (PSTN) and Internet.

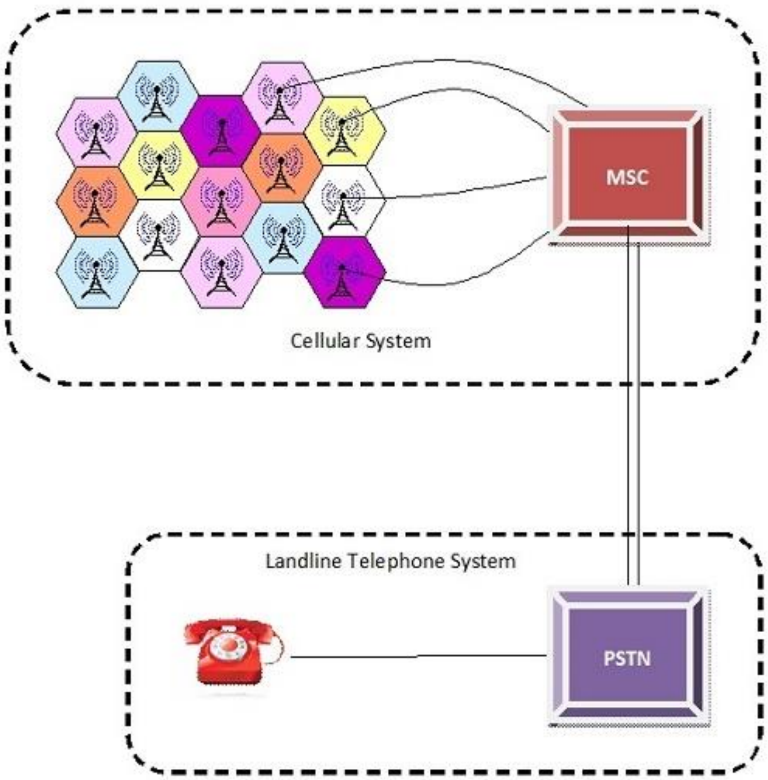


Figure 9.12: Switching Center

Chapter 10

Data Encoding & Modulation

Encoding and Modulation

Encoding is the process of converting data into a format required for a number of information processing needs, including:

- Program compiling and execution.
- Data transmission, storage and compression/decompression.
- Application data processing, such as file conversion.
- Encoding is also used to reduce the size of audio and video files.
- E.g. ASCII (American standard code for Information Interchange), MIME (Multipurpose Internet Mail Extensions).

Modulation is the process of varying one or more properties of a high frequency signal called carrier signal according with a modulating signal, which typically contains information to be transmitted. The technique of superimposing the message signal on the carrier is known as modulation. The three key parameters are: Amplitude (volume), phase (Phase) and frequency (Pitch). Modulation of a sine waveform is used to transform a baseband message into a pass band signal, for example low frequency audio signal into a radio frequency signal (RF Signal).

Encoding Techniques

The data encoding technique is divided into the following types, depending upon the type of data conversion.

- **Analog data to Analog signals** − the modulation techniques such as Amplitude Modulation, Frequency Modulation and Phase Modulation of analog signals, fall under this category.
- **Analog data to Digital signals** − this process can be termed as digitization, which is done by Pulse Code Modulation PCM. Hence, it is nothing but digital modulation. As we have already discussed, sampling and quantization are the important factors in this. Delta Modulation gives a better output than PCM.
- **Digital data to Analog signals** − the modulation techniques such as Amplitude Shift Keying ASK, Frequency Shift Keying FSK, Phase Shift Keying PSK, etc., fall under this category.
- **Digital data to Digital signals**

Digital Transmission

Data can be represented either in analog or digital form. The computers used the digital form to store the information. Therefore, the data needs to be convert into digital form so that a computer can use it.

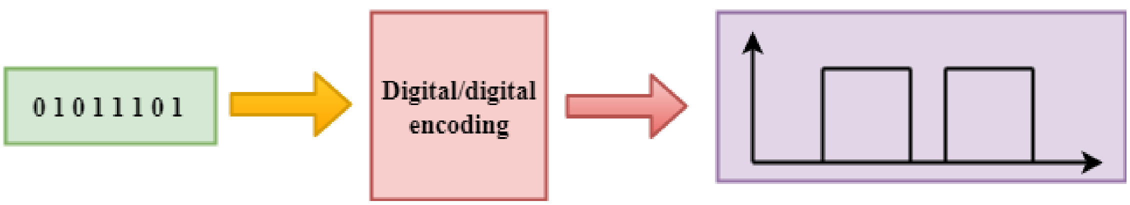
A line code is the code used for data transmission of a digital signal over a transmission line. This process of coding is chosen to avoid overlap and distortion of signal such as inter-symbol interference.

Properties of Line Coding

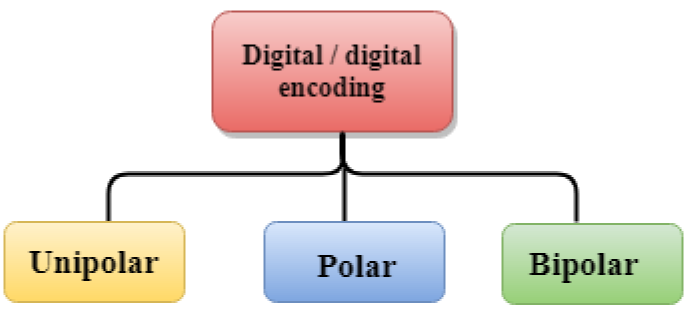
- As the coding is done to make more bits transmit on a single signal, the bandwidth used is much reduced.
- For a given bandwidth, the power is efficiently used.
- The probability of error is much reduced.
- Error detection is done and the bipolar too has a correction capability.
- Power density is much favorable.
- The timing content is adequate.
- Long strings of 1s and 0s is avoided to maintain transparency.

DIGITAL-TO-DIGITAL CONVERSION

Digital-to-digital encoding is the representation of digital information by a digital signal. When binary 1s and 0s generated by the computer are translated into a sequence of voltage pulses that can be propagated over a wire, this process is known as digital-to-digital encoding.



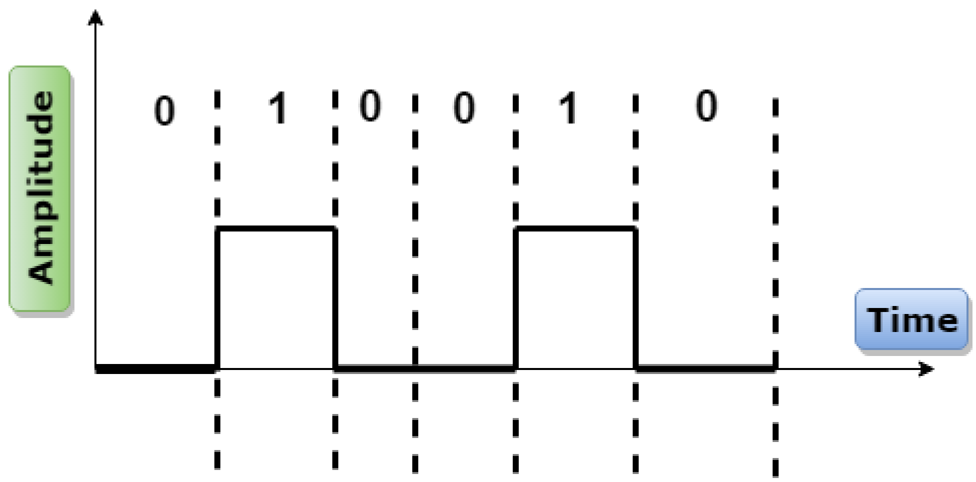
Digital-to-digital encoding divided into three categories:



i. <u>Unipolar</u>

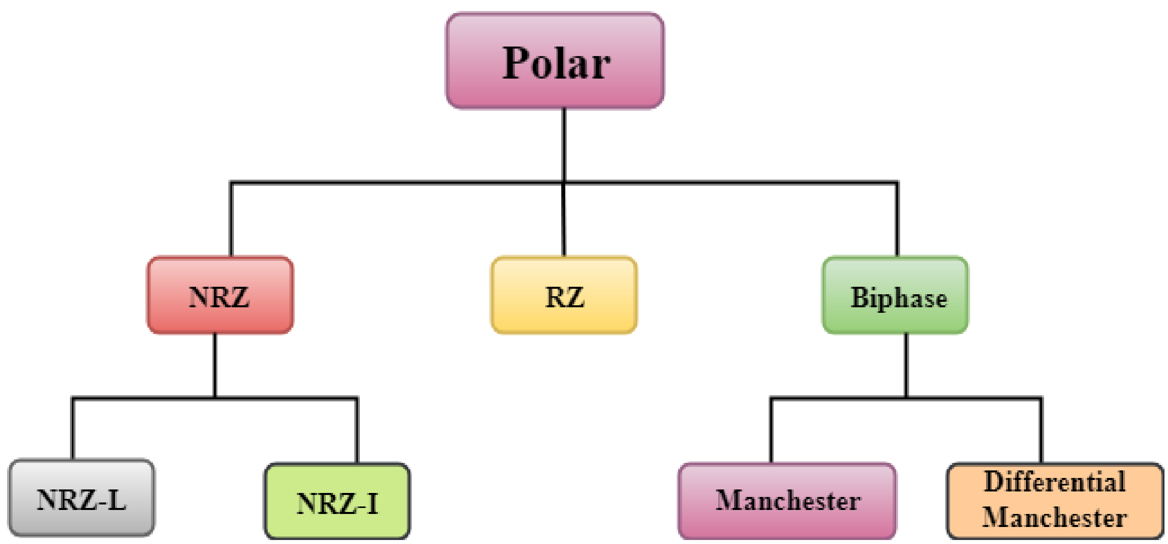
- Digital transmission system sends the voltage pulses over the medium link such as wire or cable.
- In most types of encoding, one voltage level represents 0, and another voltage level represents 1.
- The polarity of each pulse determines whether it is positive or negative.
- This type of encoding is known as Unipolar encoding as it uses only one polarity.
- In Unipolar encoding, the polarity is assigned to the 1 binary state.
- In this, 1s are represented as a positive value and 0s are represented as a zero value.

- In Unipolar Encoding, '1' is considered as a high voltage and '0' is considered as a zero voltage.
- Unipolar encoding is simpler and inexpensive to implement.
- Unipolar encoding has two problems that make this scheme less desirable:
    - DC Component
    - Synchronization



ii. Polar

- Polar encoding is an encoding scheme that uses two voltage levels: one is positive, and another is negative.
- By using two voltage levels, an average voltage level is reduced, and the DC component problem of unipolar encoding scheme is alleviated.

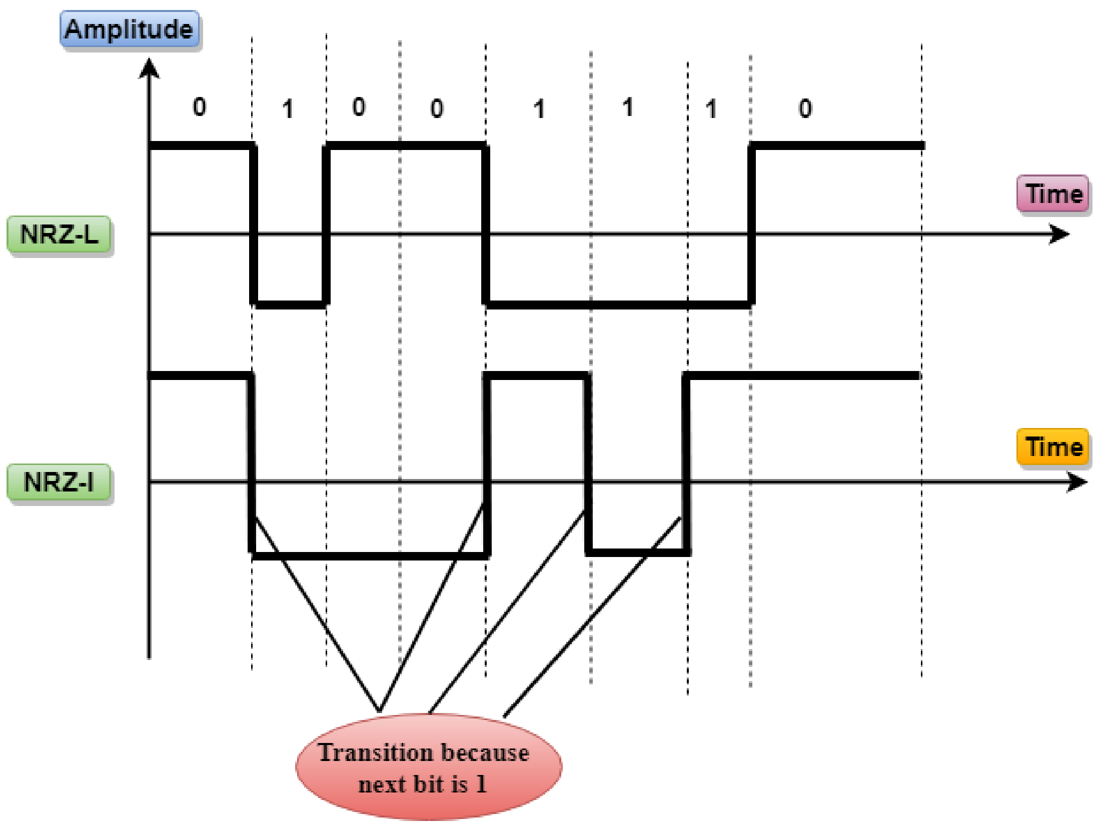


a) NRZ

- NRZ stands for Non-return zero.
- In NRZ encoding, the level of the signal can be represented either positive or negative.

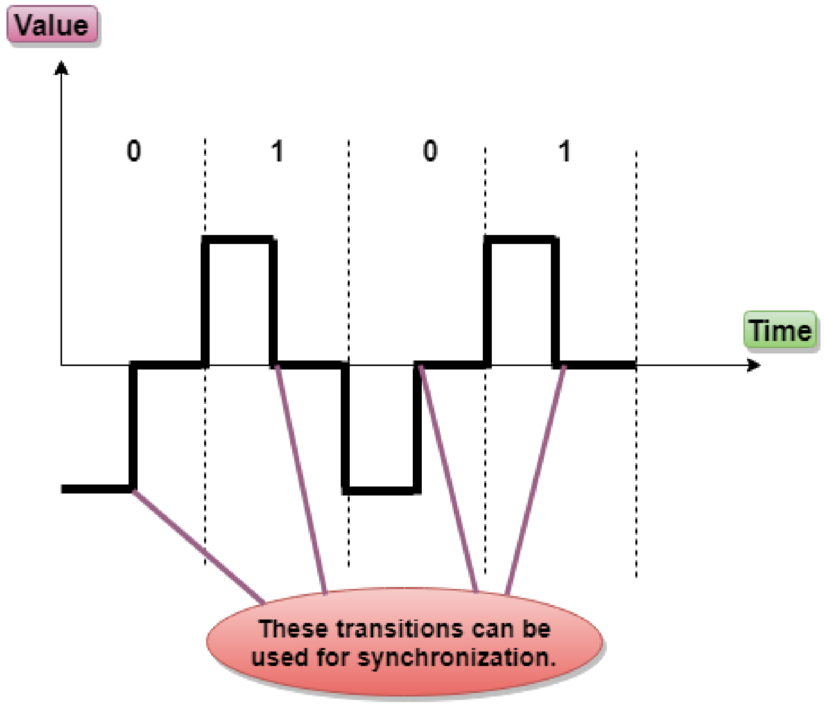
The two most common methods used in NRZ are:

- ***NRZ-L***: In NRZ-L encoding, the level of the signal depends on the type of the bit that it represents. If a bit is 0 or 1, then their voltages will be positive and negative respectively. Therefore, we can say that the level of the signal is dependent on the state of the bit.
- ***NRZ-I***: NRZ-I is an inversion of the voltage level that represents 1 bit. In the NRZ-I encoding scheme, a transition occurs between the positive and negative voltage that represents 1 bit. In this scheme, 0 bit represents no change and 1 bit represents a change in voltage level.



b) RZ

- RZ stands for Return to zero.
- There must be a signal change for each bit to achieve synchronization. However, to change with every bit, we need to have three values: positive, negative and zero.
- RZ is an encoding scheme that provides three values, positive voltage represents 1, the negative voltage represents 0, and zero voltage represents none.
- In the RZ scheme, halfway through each interval, the signal returns to zero.
- In RZ scheme, 1 bit is represented by positive-to-zero and 0 bit is represented by negative-to-zero.



- *Disadvantage* of RZ:
  It performs two signal changes to encode one bit that acquires more bandwidth.

c) Biphase

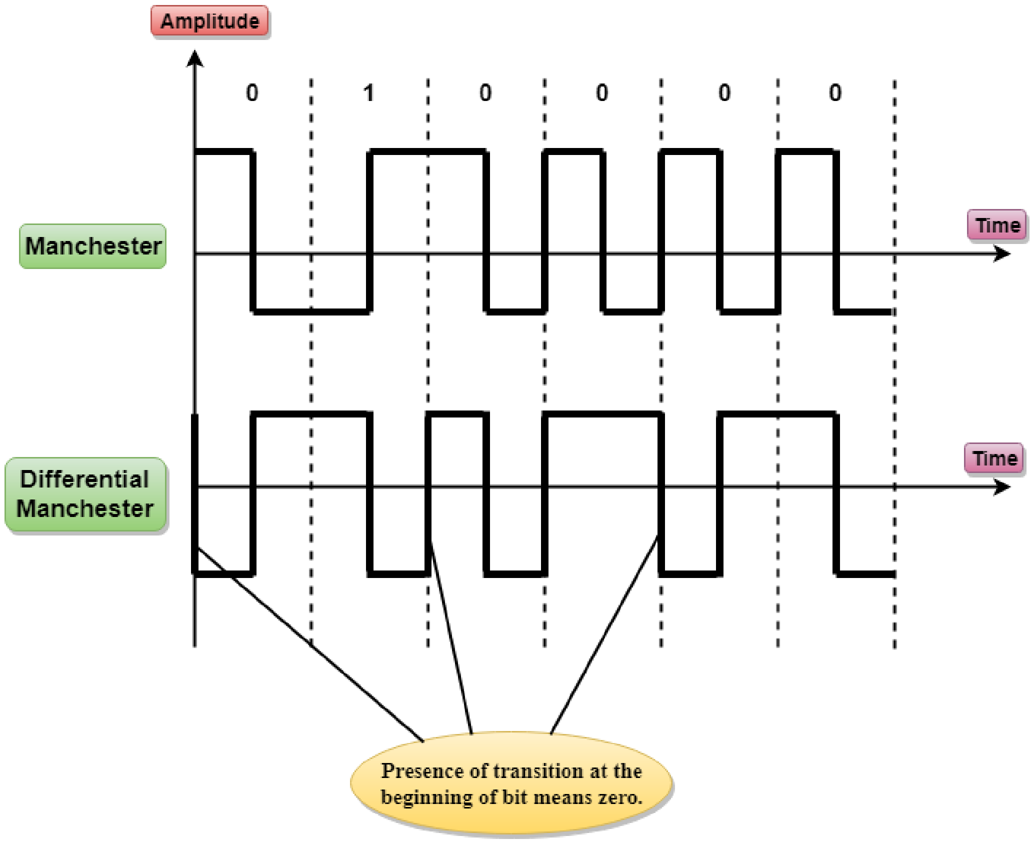
- Biphase is an encoding scheme in which signal changes at the middle of the bit interval but does not return to zero.
- Biphase encoding is implemented in two different ways:

➢ ***Manchester***

- It changes the signal at the middle of the bit interval but does not return to zero for synchronization.
- In Manchester encoding, a negative-to-positive transition represents binary 1, and positive-to-negative transition represents 0.
- Manchester has the same level of synchronization as RZ scheme except that it has two levels of amplitude.

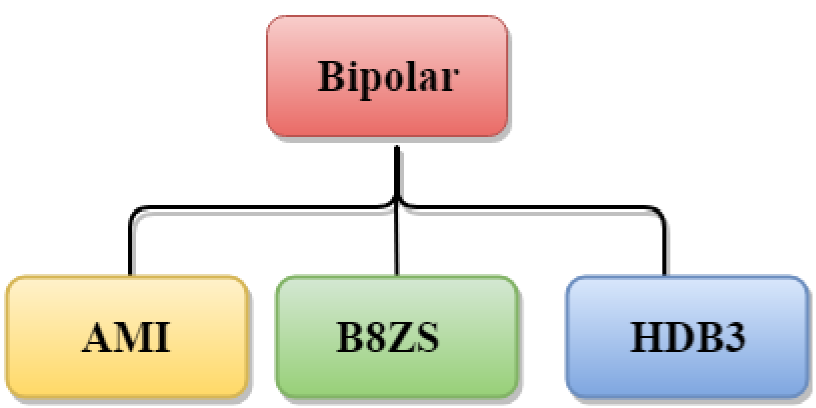
➢ ***Differential Manchester***

- It changes the signal at the middle of the bit interval for synchronization, but the presence or absence of the transition at the beginning of the interval determines the bit. A transition means binary 0 and no transition means binary 1.
- In Manchester Encoding scheme, two signal changes represent 0 and one signal change represent 1.



iii. Bipolar

- Bipolar encoding scheme represents three voltage levels: positive, negative, and zero.
- In Bipolar encoding scheme, zero level represents binary 0, and binary 1 is represented by alternating positive and negative voltages.
- If the first 1 bit is represented by positive amplitude, then the second 1 bit is represented by negative voltage, third 1 bit is represented by the positive amplitude and so on. This alternation can also occur even when the 1bits are not consecutive.
- Bipolar can be classified as:

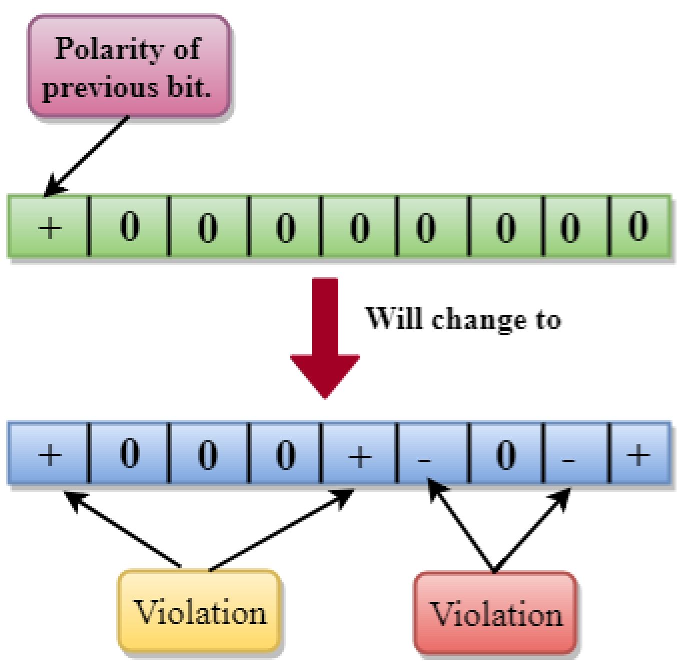


a) AMI

- AMI stands for alternate mark inversion where mark work comes from telegraphy, which means 1. So, it can be redefined as alternate 1 inversion.
- In Bipolar AMI encoding scheme, 0 bit is represented by zero level and 1 bit is represented by alternating positive and negative voltages.
- *Advantage*:
  - DC component is zero.
  - Sequence of 1s bits are synchronized.
- *Disadvantage*:
  - This encoding scheme does not ensure the synchronization of a long string of 0s bits.

b) B8ZS

- B8ZS stands for Bipolar 8-Zero Substitution.
- This technique is adopted in North America to provide synchronization of a long sequence of 0s bits.
- In most of the cases, the functionality of B8ZS is similar to the bipolar AMI, but the only difference is that it provides the synchronization when a long sequence of 0s bits occur.
- B8ZS ensures synchronization of a long string of 0s by providing force artificial signal changes called violations, within 0 string pattern.
- When eight 0 occurs, then B8ZS implements some changes in 0s string pattern based on the polarity of the previous 1 bit.
- If the polarity of the previous 1 bit is positive, the eight 0s will be encoded as zero, zero, zero, positive, negative, zero, negative, positive.

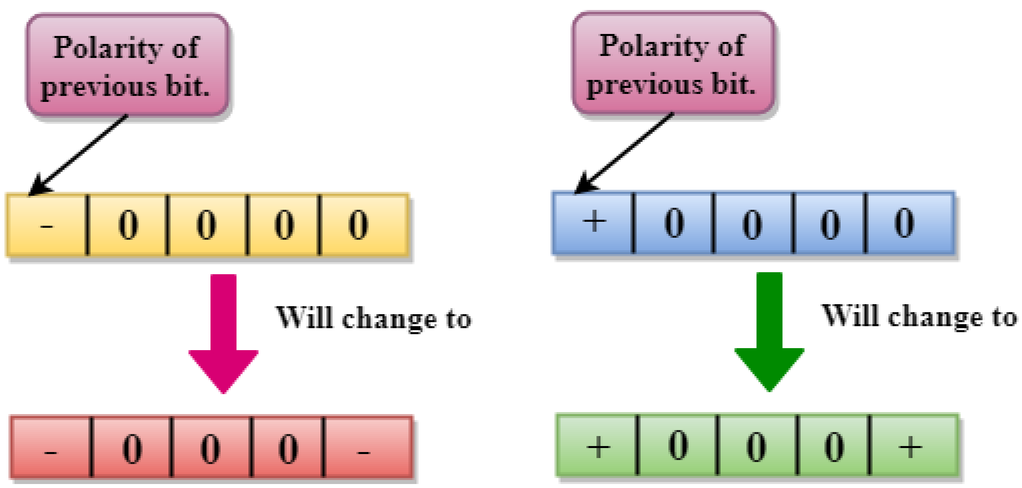


- If the polarity of previous 1 bit is negative, then the eight 0s will be encoded as zero, zero, zero, negative, positive, zero, positive, negative.

c) HDB3

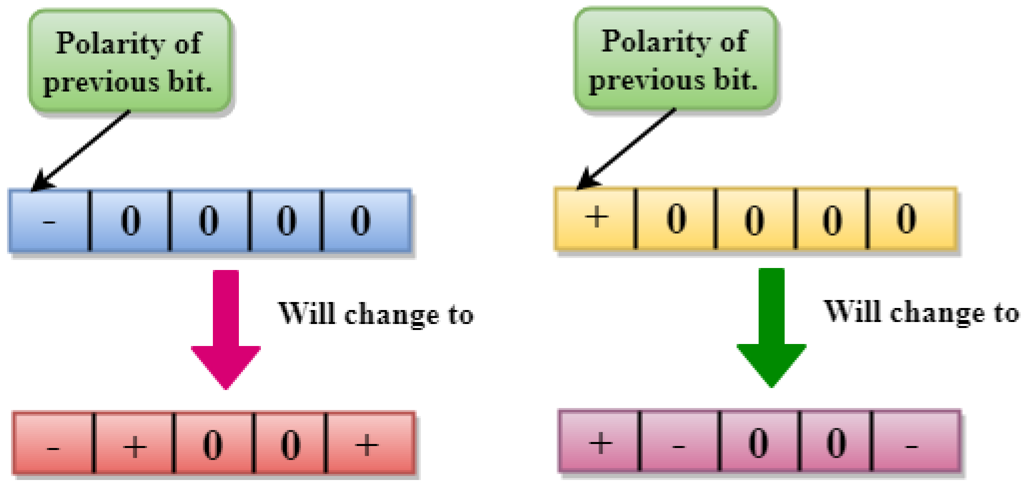
- HDB3 stands for High-Density Bipolar 3.
- HDB3 technique was first adopted in Europe and Japan.
- HDB3 technique is designed to provide the synchronization of a long sequence of 0s bits.
- In the HDB3 technique, the pattern of violation is based on the polarity of the previous bit.
- When four 0s occur, HDB3 looks at the number of 1s bits occurred since the last substitution.
- If the number of 1s bits is odd, then the violation is made on the fourth consecutive of 0. If the polarity of the previous bit is positive, then the violation is positive. If the polarity of the previous bit is negative, then the violation is negative.

- ***If the number of 1s bits since the last substitution is odd.***



If the number of 1s bits is even, then the violation is made on the place of the first and fourth consecutive 0s. If the polarity of the previous bit is positive, then violations are negative, and if the polarity of the previous bit is negative, then violations are positive.

- ***If the number of 1s bits since the last substitution is even.***



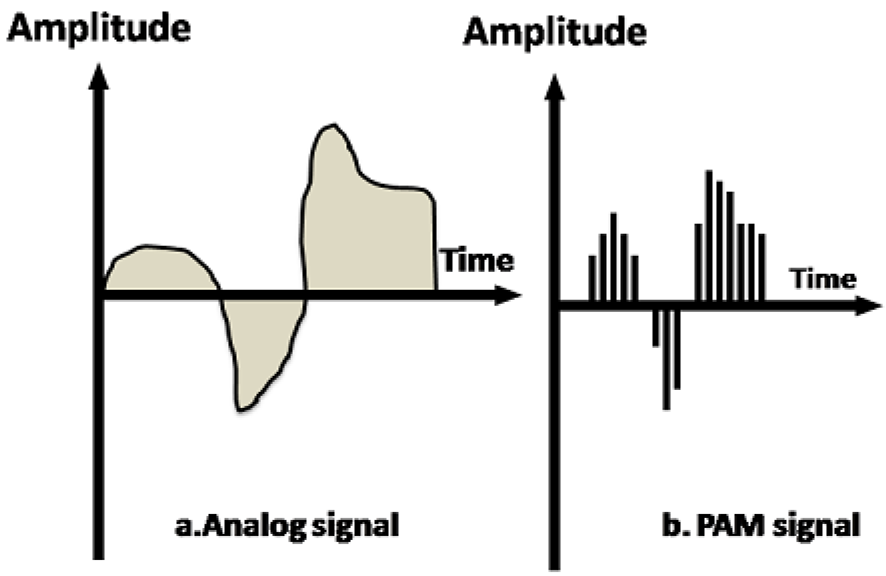
ANALOG-TO-DIGITAL CONVERSION

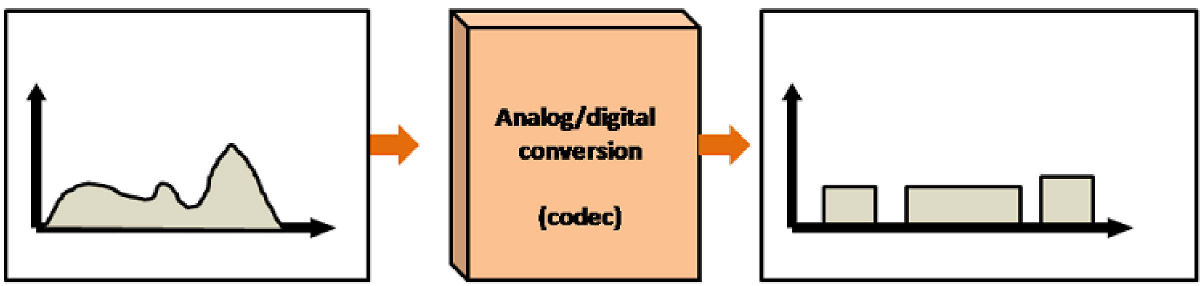
- When an analog signal is digitalized, this is called an analog-to-digital conversion.
- Suppose human sends a voice in the form of an analog signal, we need to digitalize the analog signal which is less prone to noise. It requires a reduction in the number of values in an analog message so that they can be represented in the digital stream.
- In analog-to-digital conversion, the information contained in a continuous wave form is converted in digital pulses.

Techniques for Analog-To-Digital Conversion

i. PAM

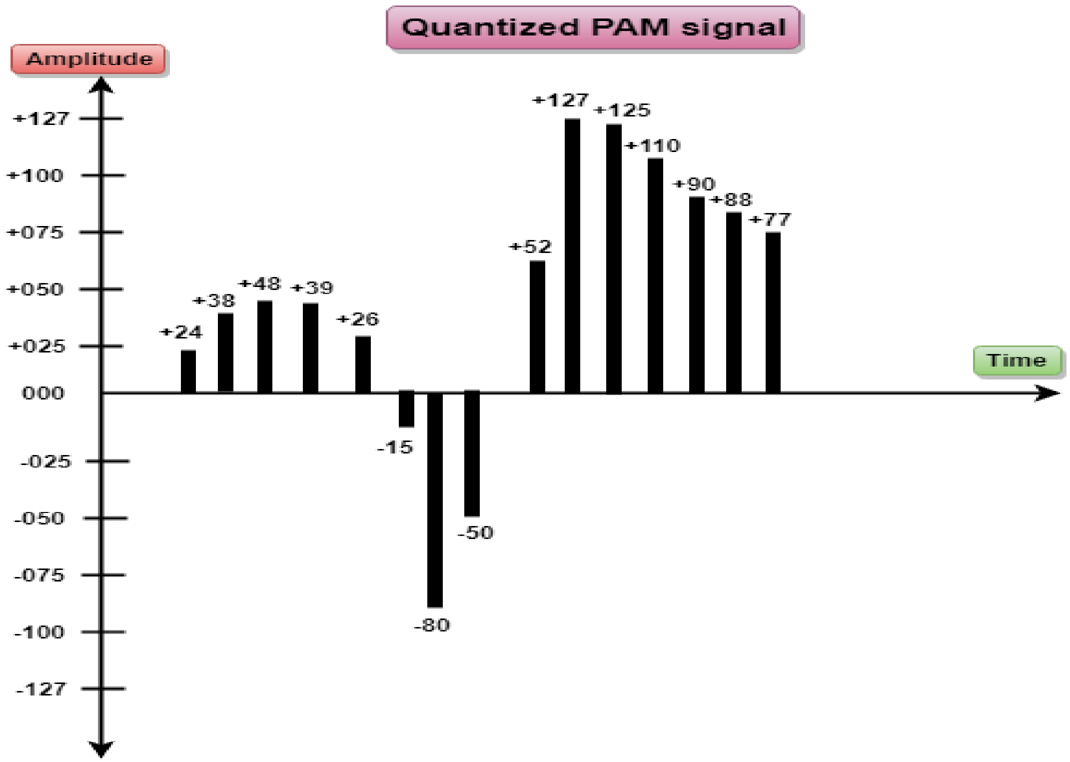
- PAM stands for pulse amplitude modulation.
- PAM is a technique used in analog-to-digital conversion.
- PAM technique takes an analog signal, samples it, and generates a series of digital pulses based on the result of sampling where sampling means measuring the amplitude of a signal at equal intervals.
- PAM technique is not useful in data communication as it translates the original wave form into pulses, but these pulses are not digital. To make them digital, PAM technique is modified to PCM technique.





ii. PCM

- PCM stands for Pulse Code Modulation.
- PCM technique used to modify the pulses created by PAM to form a digital signal. To achieve this, PCM quantizes PAM pulses. Quantization is a process of assigning integral values in a specific range to sampled instances.
- PCM is made of four separate processes: PAM, quantization, binary encoding, and digital-to-digital encoding.



Digital to analog

To send the digital data over an analog media, it needs to be converted into analog signal. There can be two cases according to data formatting.

Band-pass: The filters are used to filter and pass frequencies of interest. A bandpass is a band of frequencies, which can pass the filter.

Low-pass: Low-pass is a filter that passes low frequencies signals.
When digital data is converted into a bandpass analog signal, it is called digital-to-analog conversion. When low-pass analog signal is converted into bandpass analog signal, it is called analog-to-analog conversion.

Digital-to-Analog Conversion

When data from one computer is sent to another via some analog carrier, it is first converted into analog signals. Analog signals are modified to reflect digital data.
An analog signal is characterized by its amplitude, frequency, and phase. There are three kinds of digital-to-analog conversions:

i. **Amplitude Shift Keying**

In this conversion technique, the amplitude of analog carrier signal is modified to reflect binary data.

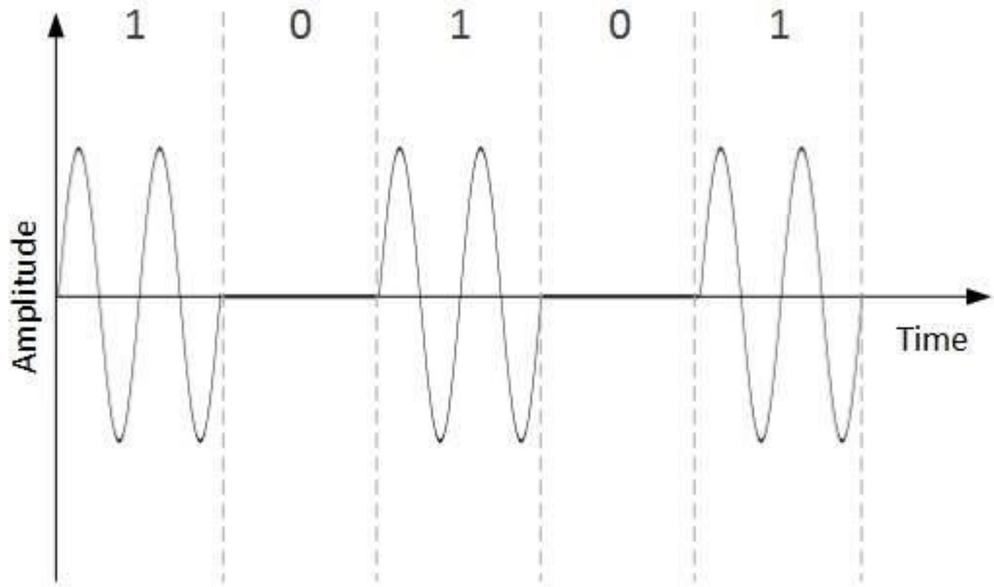


Figure 10.43: Amplitude Shift Keying

When binary data represents digit 1, the amplitude is held; otherwise, it is set to 0. Both frequency and phase remain same as in the original carrier signal.

ii. Frequency Shift Keying

In this conversion technique, the frequency of the analog carrier signal is modified to reflect binary data.

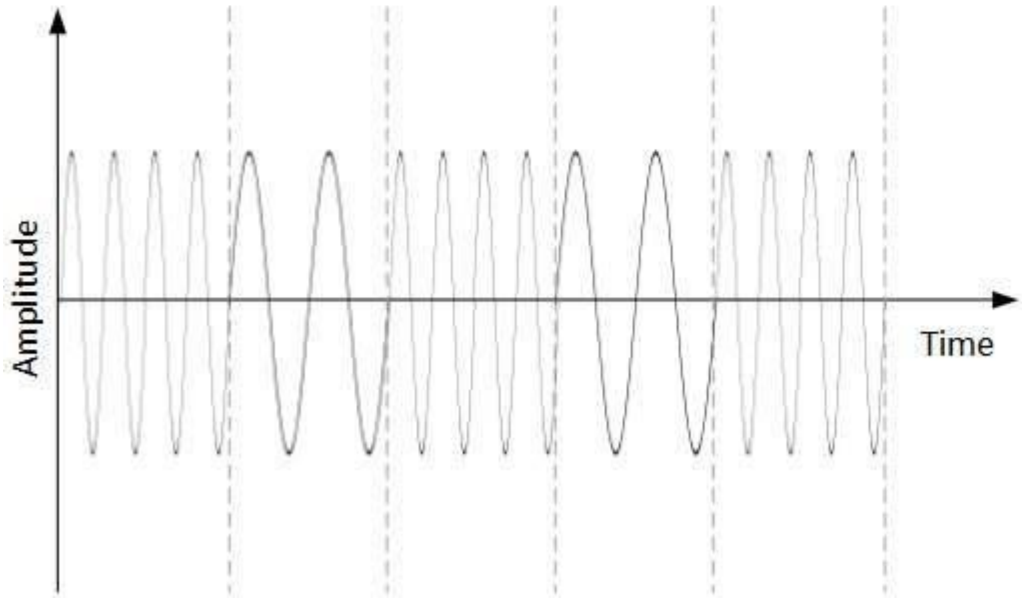


Figure 10.44: Frequency Shift Keying

This technique uses two frequencies, f1 and f2. One of them, for example f1, is chosen to represent binary digit 1 and the other one is used to represent binary digit 0. Both amplitude and phase of the carrier wave kept intact.

iii. Phase Shift Keying

In this conversion scheme, the phase of the original carrier signal is altered to reflect the binary data.

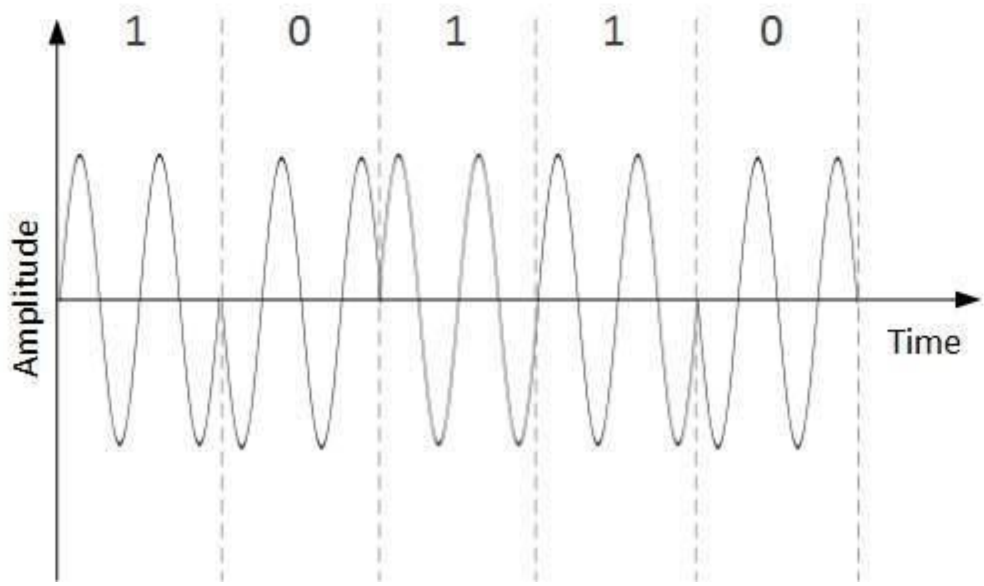


Figure 10.45: Phase Shift Keying

When a new binary symbol is encountered, the phase of the signal is altered. Amplitude and frequency of the original carrier signal is kept intact.

iv. Quadrature Phase Shift Keying

QPSK alters the phase to reflect two binary digits at once. This is done in two different phases. The main stream of binary data is divided equally into two sub-streams. The serial data is converted in to parallel in both sub-streams and then each stream is converted to digital signal using NRZ technique. Later, both the digital signals are merged together.

Analog-to-Analog Conversion

Analog-to-analog conversion, or modulation, is the representation of analog information by an analog signal. A characteristic of carrier wave is varied according to the instantaneous amplitude of the modulating signal by virtue of a process. This modulation is generally needed when a bandpass channel is required. Bandpass is a range of frequencies, which are transmitted through a bandpass filter which is a filter allowing specific frequencies to pass preventing signals at unwanted frequencies.

Analog-to-Analog conversion can be done in three ways:

1. Amplitude Modulation
2. Frequency Modulation
3. Phase Modulation

1. **Amplitude Modulation**

Amplitude modulation is a process by which the wave signal is transmitted by modulating the amplitude of the signal. It is often called as AM and is commonly used in transmitting a piece of information through a radio carrier wave. Amplitude modulation is mostly used in the form of electronic communication.

Currently, this technique is used in many areas of communication such as in portable two-way radios, citizens band radio, VHF aircraft radio and in modems for computers. Amplitude modulation is also used to mention the mediumwave AM radio broadcasting.

What is Amplitude Modulation?

Amplitude modulation or just AM is one of the earliest modulation methods that is used in transmitting information over the radio. This technique was devised in the 20th century at a time when Landell de Moura and Reginald Fessenden were conducting experiments using a radiotelephone in the 1900s. After successful attempts, the modulation technique was established and used in electronic communication.

In general, amplitude modulation definition is given as a type of modulation where the amplitude of the carrier wave is varied in some proportion with respect to the modulating data or the signal.

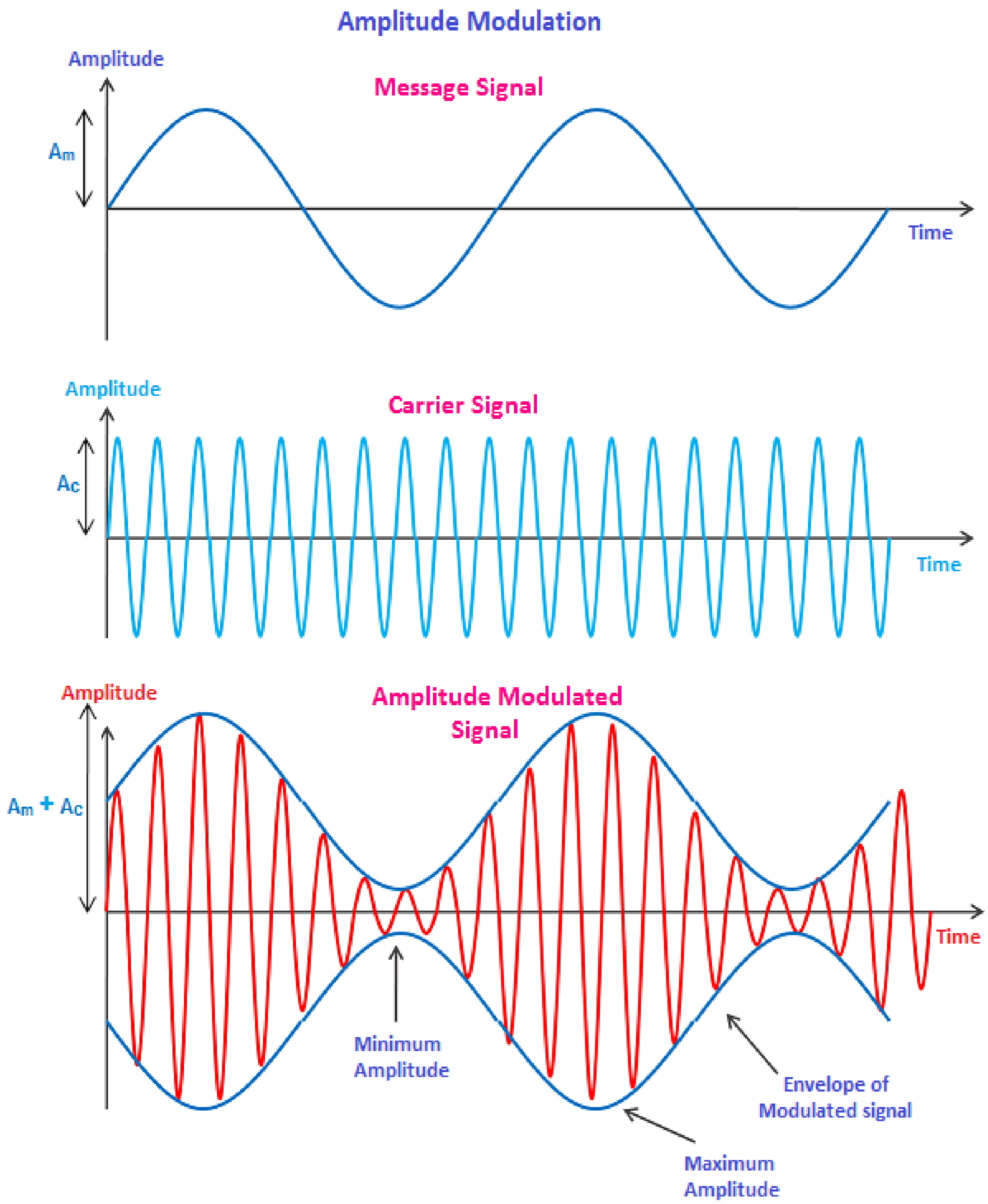
As for the mechanism, when amplitude modulation is used there is a variation in the amplitude of the carrier. Here, the voltage or the power level of the information signal changes the amplitude of the carrier. In AM, the carrier does not vary in amplitude. However, the modulating data is in the form of signal components consisting of frequencies either higher or lower than that of the carrier. The signal components are known as sidebands and the sideband power is responsible for the variations in the overall amplitude of the signal.

The AM technique is totally different from frequency modulation and phase modulation where the frequency of the carrier signal is varied in the first case and in the second one the phase is varied respectively.

Types of Amplitude Modulation

There are three main types of amplitude modulation. They are;

- Double sideband-suppressed carrier modulation (DSB-SC).
- Single Sideband Modulation (SSB).
- Vestigial Sideband Modulation (VSB).



Carrier wave $c(t) = A_c \, sin \, \omega_c t$

Modulation signal $m(t) = A_m \sin \omega_m t$

Amplitude modulated Wave $m(t) = (A_c + A_m \sin \omega_m t) \, sin \, \omega_c t$

Modulation index $(\mu) = \frac{A_m}{A_c} = \frac{Amplitude \ of \ modulating \ signal}{Amplitude \ of \ Carrier \ wave}$

Bandwidth $(w) = 2f_m$

Advantages of Amplitude Modulation

1. **Few components needed:** At the receiver side, the original signal is extracted (demodulated) using a circuit consisting of very few components.

2. **Low cost:** The components used in amplitude modulation is very cheap. So the AM transmitter and AM receiver build at low cost.

3. **It is simple to implement.**

4. **Long distance communication:** Amplitude modulated waves can travel a longer distance.

Disadvantages of Amplitude Modulation

1. **Amplitude modulation is inefficient in terms of its power usage:** As we know that the message signal contains information whereas the carrier signal does not contain any information. In amplitude modulation, most of the power is concentrated in the carrier signal which contains no information. At the receiver side, the power consumed by the carrier wave is wasted.

2. **It requires high bandwidth:** The amplitude modulation is not efficient in terms of its use of bandwidth. It requires a bandwidth equal to twice that of the highest audio signal frequency.

3. This type of transmission can be easily affected by the external radiation.

4. This type of transmission is also affected by the man-made noises or radiations like waves from other antennas or channels.

5. Amplitude modulation (AM) cannot be used for transmitting music as done by frequency modulation (FM).

6. Amplitude modulation cannot be used for transmission of sensitive information like in the army, where interpretation or loss or disruption during transmission is not an option.

Applications of Amplitude Modulation

1. **Air band radio:** The amplitude modulation is extensively used in aerospace industry. The VHF (Very High Frequency) transmissions made by the airborne equipment still use amplitude modulation. The radio contact between ground to ground and also ground to air use amplitude modulated (AM) signals.

2. **Broadcast transmission:** Amplitude modulation (AM) is still widely used for broadcasting either short or medium or long wave bands.

3. **Quadrature amplitude modulation:** Amplitude modulation is used in the transmission of data of almost everything, from short-range transmission such as wifi to cellular communications. Quadrature amplitude modulation is formed by mixing two carriers that are out of phase by 90°.

4. **Single sideband:** The amplitude modulation (AM) in the form of single sideband is still used for HF (High Frequency) radio links.

2. <u>Frequency Modulation</u>

Frequency modulation is a type of modulation where the frequency of the carrier signal varies as per amplitude variations of the message signal.

or

Frequency modulation is a type of modulation where the information (message signal) is transmitted over a carrier wave by varying its frequency in accordance with the amplitude of the message signal.

or

Frequency modulation is the process of superimposing the message signal onto the carrier signal and the resulting wave with variable frequency is called a frequency modulated wave.

or

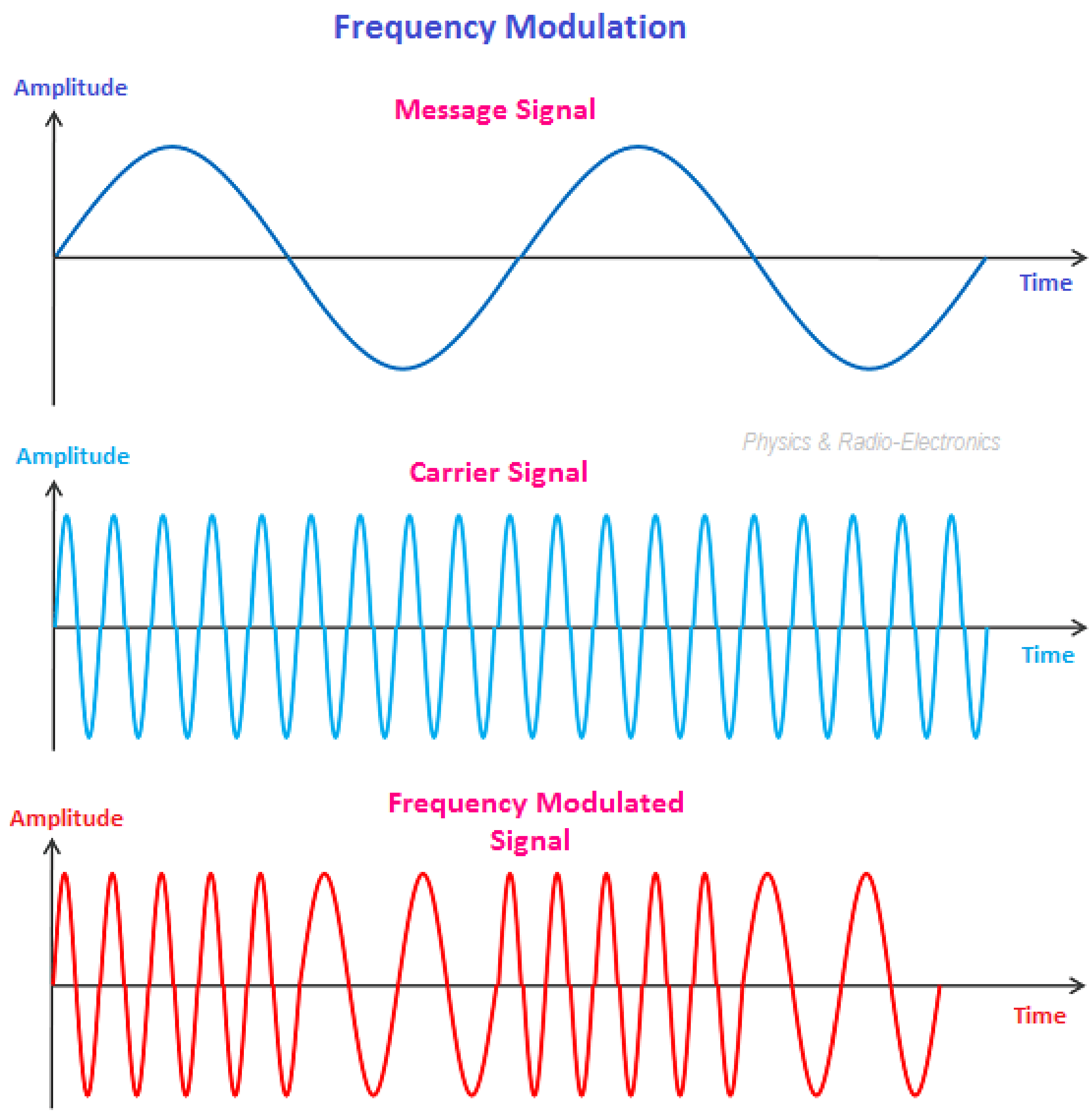
Frequency modulation is the process of transmitting information over a carrier wave by varying its frequency in accordance with the amplitude of the message signal.

Frequency Modulation Diagram

In frequency modulation, the frequency of the carrier signal is varied whereas the amplitude of the carrier signal remains constant. Frequency modulation is also simply referred to as FM.

The below figures show the frequency modulation.

The first figure shows the message signal or modulating signal which contains information, the second figure shows the high frequency carrier wave which contains no information, and the last figure shows the resultant frequency modulated wave.



The third figure shows that the frequency of both the positive and negative half cycles of the carrier wave is varied in accordance with the instant amplitude of the modulating signal (message signal).

In frequency modulation, the amount of change in frequency of the carrier signal is determined by the amplitude of the message signal. For example, let's assume that the carrier signal has a frequency deviation of ± 4 kHz. In such case, the carrier signal will move up and down by 4 kHz. The frequency of the modulated wave is high when the message signal reaches its maximum amplitude.

The carrier wave does not contain any information so even if we change the frequency of the carrier wave, there will be no information loss. However, if we change the frequency of the modulating signal, some amount of information loss will occur because the modulating signal contains the information. So the frequency of the modulating signal should not be changed.

Amplitude modulation was the first modulation technique developed to transmit voice signals by using radio waves. Frequency modulation was developed after amplitude modulation. The main advantage of frequency modulation over amplitude modulation is that it is more resistant to additive noise than amplitude modulation.

As we know that the message signal contains information. However, we cannot transmit message signals to very large distances because of its low signal strength. Hence we use a high signal strength carrier wave to transmit information over very large distances. In simple words, the message signal takes the help of a carrier signal to transmit information over very large distances.

The frequency modulation technique is widely used for FM broadcasting. It is also used in radar, telemetry, seismic prospecting, video transmission systems, music synthesis, two-way radio systems, and magnetic tape recording systems.

The main advantage of frequency modulation (FM) in radio transmission is that it has a large signal-to-noise ratio (ratio of signal power to the noise power) and therefore rejects radio frequency interference better than the amplitude modulation (AM). Hence, most music is broadcast over FM radio.

The frequency modulation and phase modulation belongs to the family of angle modulation. Frequency and phase modulation are dependent on each other. When the frequency of the carrier signal varies, the phase of the carrier signal also varies. Similarly, when the phase of the carrier signal varies, the frequency of the carrier signal also varies. However, if frequency is varied directly, then it is called frequency modulation. And if phase is varied directly, then it is called phase modulation.

Modulation Index of Frequency Modulation

Frequency modulation index describes how the frequency of the carrier signal and amplitude of message signal affects the frequency of the frequency modulated (FM) signal.

or

Frequency modulation index is defined as the ratio of maximum frequency deviation of the carrier signal to the frequency of the message signal.

$$\text{Modulation Index } (M_i) = \frac{\Delta f}{f_m}$$

Where, Δf = Maximum frequency deviation of the carrier signal

F_m = Frequency of the message signal

Frequency Modulation Deviation Ratio

Frequency modulation deviation ratio is defined as the ratio of the maximum carrier frequency deviation to the highest message signal frequency.

$$D = \frac{\text{Maximum carrier frequency deviation}}{\text{Maximum message signal frequency}}$$

Types of frequency modulation

Normally an unmodulated carrier signal contains a single frequency. However, when the carrier signal is mixed with the message signal, two additional frequencies are created for the frequency modulated signal. These additional frequencies are known as sidebands. The highest frequency of the carrier signal is known as upper sideband and the lowest frequency of the carrier signal is known as lower sideband.

The bandwidth of the modulated signal can be obtained by taking the difference between the highest and lowest frequencies of the modulated signal. The bandwidth of an FM signal is depends on the frequency deviation. When the frequency deviation is high, the bandwidth required will be large; and when the frequency deviation is low, the bandwidth required is low. Depending on the bandwidth requirement, frequency modulation can be divided into two types: Wideband FM and Narrowband FM.

Wideband FM

In wideband FM, the modulation index normally exceeds 1.

The frequency deviation is very high in wideband FM. However, the maximum permissible frequency deviation is 75 KHz.

The wideband FM has infinite bandwidth.

The noise will be better suppressed in wideband FM because of its large frequency deviation.

Wideband FM occupies up to 15 times the bandwidth of the narrowband FM.

Because of its high quality transmission, wideband FM is used in entertainment broadcasting.

Narrowband FM

The modulation index in narrowband FM is nearly unity.

The frequency deviation is very low in narrowband FM. The maximum frequency deviation is 5 kHz.

Narrowband FM is frequently used for short distance communications.

Because of its low quality transmission, narrowband FM is used for mobile communications.

Advantages of Frequency modulation

i. All the power transmitted in frequency modulation is useful whereas in amplitude modulation, most of the power is in carrier (which is useless).
ii. Adjacent channel interference does not takes place in frequency modulation.
iii. High signal to noise ratio (S/N). In simple words, it has less noise.

Drawbacks of Frequency modulation

It requires wider bandwidth than amplitude modulation.

Applications of frequency modulation

1. FM broadcasting
2. Radar
3. Magnetic tape recording systems
4. Telemetry
5. Two-way radio systems
6. Music synthesis
7. Seismic prospecting
8. Video transmission systems

3. **<u>Phase Modulation</u>**

Phase modulation is a type of modulation where the phase of the carrier signal varies as per amplitude variations of the message signal.

or

Phase modulation is a type of angle modulation in which the total phase angle of the carrier signal is varied in accordance with the amplitude of the message signal.

or

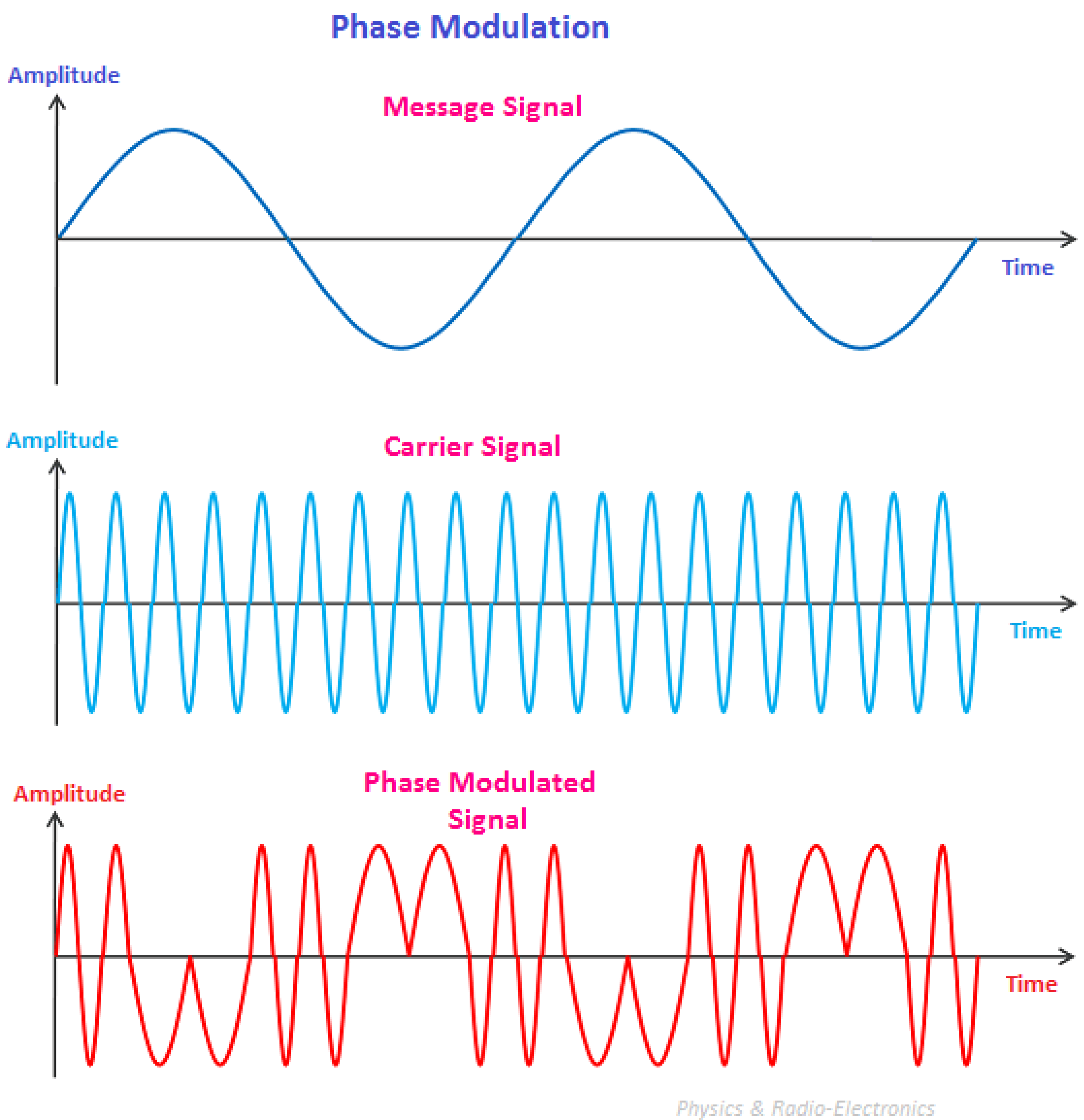
Phase modulation is the process of transmitting information over a carrier wave by varying its phase in accordance with the amplitude of the message signal.

Phase Modulation Diagram

In phase modulation, the phase of the carrier signal is varied whereas the amplitude of the carrier signal remains constant. Phase modulation is also referred to as PM.

The below three figures show the phase modulation.

The first figure shows the low frequency modulating signal or message signal which contains useful information, the second figure shows the high frequency carrier wave which does not contain any information, and the last figure shows the resultant phase modulated signal.



The third figure shows that the phase of both the positive and negative half cycles of the carrier signal varies as per amplitude variations of the modulating signal. During the positive half cycle, the carrier signal phase shifts in one direction, whereas during the negative half cycle, the carrier signal phase shifts in the opposite direction.

In phase modulation, the phase deviation is directly proportional to the amplitude of message signal.

The noise immunity of the phase modulation is better than amplitude modulation. However, the noise immunity of the phase modulation is not as good as frequency modulation.

The signal-to-noise (SNR) ratio of the phase modulation is better than amplitude modulation. However, the signal-to-noise (SNR) of the phase modulation is not as good as frequency modulation.

The modulation index of phase modulation is directly proportional to the message signal.

Phase modulation is primarily used for some mobile radio services.

The phase modulation and frequency modulation are closely related to each other. In both phase and frequency modulation, the total phase angle of the modulated signal varies.
In practice, the phase modulation and frequency modulation are dependent on each other. When the phase of the carrier signal varies, the frequency of the carrier signal also varies. Similarly, when the frequency of the carrier signal varies, the phase of the carrier signal also varies. However, the phase modulation and frequency modulation are not directly proportional to each other.
In frequency modulation, the total phase angle of a carrier wave changes for a short period due to the change in frequency of the carrier wave. In phase modulation, the frequency of a carrier wave changes for a short period due to the change in phase of a carrier wave.