UNIT-1: RADAR & NAVIGATION AIDS

BASIC RADAR-

RADAR stands for Radio Detection and Ranging System. It is basically an electromagnetic system used to detect the location and distance of an object from the point where the RADAR is placed. It works by radiating energy into space and monitoring the echo or reflected signal from the objects. It operates in the UHF and microwave range.

A radar is an electromagnetic sensor, used to notice, track, locate, and identify different objects which are at certain distances. The working of radar is, it transmits electromagnetic energy in the direction of targets to observe the echoes and returns from them. Here the targets are nothing but ships, aircraft, astronomical bodies, automotive vehicles, spacecraft, rain, birds, insects, etc. Instead of noticing the target's location and velocity, it also obtains their shape and size sometimes.

The main objective of radar as compared with infrared and optical sensing devices is to discover faraway targets under difficult climate conditions & determines their distance, range, through precision. Radar has its own transmitter which is known as a source of illumination for placing targets. Generally, it works in the microwave area of the electromagnetic spectrum that is calculated in hertz when frequencies extend from 400 MHz to 40 GHz.

ADVANTAGES-

- RADAR signal can penetrate mediums such as clouds, fogs, mist, and snow.
- It can give the exact position, velocity and distance of an object.
- It can tell the difference between stationary and moving targets.
- RADAR signals do not require a medium of transportation.
- RADAR signals can target several objects simultaneously.
- It is wireless and does not rely on wire connectivity.
- It is cheaper as compared to other systems.
- It covers a wider geographical area.
- It allows for repetitive coverage.
- It is fast if the area is not too large.
- It can get data from some of the remote areas of the planet.

APPLICATIONS-

The applications of radar include the following.

Military Applications

It has 3 major applications in the Military:

- In air defense, it is used for target detection, target recognition, and weapon control (directing the weapon to the tracked targets).
- In a missile system to guide the weapon.
- Identifying enemy locations on the map.

Air Traffic Control

It has 3 major applications in Air Traffic control:

- To control air traffic near airports. The Air Surveillance RADAR is used to detect and display the aircraft's position in the airport terminals.
- To guide the aircraft to land in bad weather using Precision Approach RADAR.
- To scan the airport surface for aircraft and ground vehicle positions

Remote Sensing

It can be used for observing whether or observing planetary positions and monitoring sea ice to ensure a smooth route for ships.

Ground Traffic Control

It can also be used by traffic police to determine the speed of the vehicle, controlling the movement of vehicles by giving warnings about the presence of other vehicles or any other obstacles behind them.

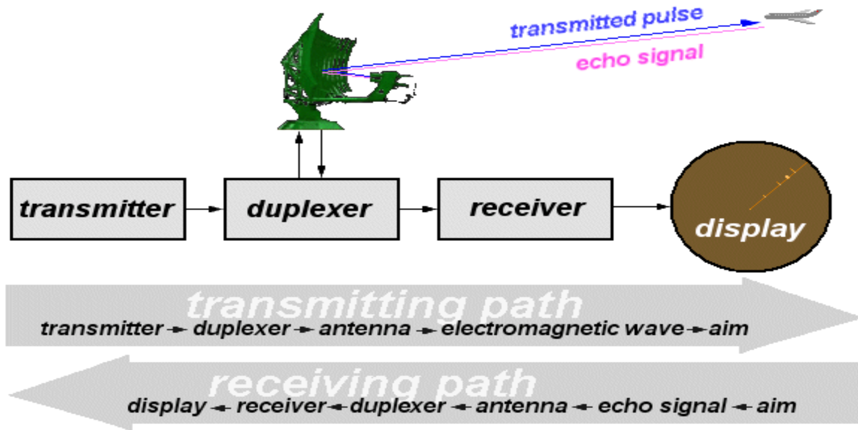
Space

It has 3 major applications

- To guide the space vehicle for a safe landing on the moon
- To observe the planetary systems
- To detect and track satellites
- To monitor the meteors

WORKING PRINCIPLE OF SIMPLE RADAR SYSTEM-

The radar antenna radiates microwave signal towards target, which is then reflected and picked up by a receiving device. The electrical signal picked up by the receiving antenna is called echo or return. The radar signal is generated by a powerful transmitter and received by a highly sensitive receiver. Radar signals can be displayed on the display.



- **Transmitter**
  The radar transmitter produces the short duration high-power radio frequency pulses of energy that are into space by the antenna.
- **Duplexer**
  The duplexer alternately switches the antenna between the transmitter and receiver so that only one antenna need be used. This switching is necessary because the high-power pulses of the transmitter would destroy the receiver if energy were allowed to enter the receiver.
- **Receiver**
  The receivers amplify and demodulate the received RF-signals. The receiver provides video signals on the output.
- **Radar Antenna**
  The Antenna transfers the transmitter energy to signals in space with the required distribution and efficiency. This process is applied in an identical way on reception.
- **Indicator**
  The indicator should present to the observer a continuous, easily understandable, graphic picture of the relative position of radar targets. The radar screen (in this case a PPI-scope) displays the output produced from the echo signals.

TYPES-

There are two types of radar system

1-Pulsed radar

2-Continuous wave radar

The pulse radar transmits short rectangular pulses & the continuous wave radar transmits continuous sinusoidal EM waves.

<u>Pulse Radar</u>

The Radar, which operates with pulse signal is called the **Pulse Radar**. Pulse Radars can be classified into the following two types based on the type of the target it detects.

- Basic Pulse Radar

- Moving Target Indication Radar

Basic Pulse Radar

The Radar, which operates with pulse signal for detecting stationary targets, is called the **Basic Pulse Radar** or simply, Pulse Radar. It uses single Antenna for both transmitting and receiving signals with the help of Duplexer.

Antenna will transmit a pulse signal at every clock pulse. The duration between the two clock pulses should be chosen in such a way that the echo signal corresponding to the present clock pulse should be received before the next clock pulse.

Moving Target Indication Radar

The Radar, which operates with pulse signal for detecting non-stationary targets, is called Moving Target Indication Radar or simply, **MTI Radar**. It uses single Antenna for both transmission and reception of signals with the help of Duplexer.

MTI Radar uses the principle of **Doppler Effect** for distinguishing the non-stationary targets from stationary objects.

Continuous Wave Radar

The Radar, which operates with continuous signal or wave is called **Continuous Wave Radar**. They use Doppler Effect for detecting non-stationary targets. Continuous Wave Radars can be classified into the following two types.

- Unmodulated Continuous Wave Radar
- Frequency Modulated Continuous Wave Radar

Unmodulated Continuous Wave Radar

The Radar, which operates with continuous signal (wave) for detecting non-stationary targets is called Unmodulated Continuous Wave Radar or simply, **CW Radar**. It is also called CW Doppler Radar.

This Radar requires two Antennas. Of these two antennas, one Antenna is used for transmitting the signal and the other Antenna is used for receiving the signal. It measures only the speed of the target but not the distance of the target from the Radar.

Frequency Modulated Continuous Wave Radar

If CW Doppler Radar uses the Frequency Modulation, then that Radar is called the Frequency Modulated Continuous Wave **(FMCW)** Radar or FMCW Doppler Radar. It is also called Continuous Wave Frequency Modulated Radar or CWFM Radar.

This Radar requires two Antennas. Among which, one Antenna is used for transmitting the signal and the other Antenna is used for receiving the signal. It measures not only the speed of the target but also the distance of the target from the Radar.

RADAR RANGE EQUATION

The standard form of Radar range equation is also called as simple form of Radar range equation. Now, let us derive the standard form of Radar range equation.

We know that power density is nothing but the ratio of power and area. So, the power density (P_{di}) at a distance, R from the Radar can be mathematically represented as –

$$P_{di} = P_t \ / \ 4\pi R^2 \qquad \text{(equation 1)}$$

Where,

Pt is the amount of power transmitted by the Radar transmitter

The above power density is valid for an isotropic Antenna. In general, Radars use directional Antennas. Therefore, the power density, P_{dd} due to directional Antenna will be

$$P_{dd}=P_t G \ / \ 4\pi R^2 \qquad \text{(equation 2)}$$

Target radiates the power in different directions from the received input power. The amount of power, which is reflected back towards the Radar depends on its cross section. So, the power density P_{de} of echo signal at Radar can be mathematically represented as –

$$P_{de}=P_{dd}(S \ / \ 4\pi R^2) \qquad \text{(equation 3)}$$

Substitute, Equation 2 in Equation 3.

$$P_{de} = (P_t \ / G 4\pi R^2)(S \ / \ 4\pi R^2) \qquad \text{(Equation4)}$$

The amount of **power,** P_r **received** by the Radar depends on the effective aperture, A_e of the receiving Antenna.

$$P_r = P_{de} A_e \qquad \text{(Equation5)}$$

Substitute, Equation 4 in Equation 5.

$P_r = (P_t G \ / \ 4\pi R^2)(S \ / \ 4\pi R^2) \ A_e$

$\Rightarrow P_r = P_t GSA_e \ / \ (4\pi)^2 R^4$

$\Rightarrow R^4 = P_t GSA_e \ / \ (4\pi)^2 P_r$

$$\Rightarrow R = \left[\frac{P_t GSA_e}{(4\pi)^2 P_r}\right]^{1/4} \qquad \text{(Equation 6)}$$

Standard Form of Radar Range Equation

If the echo signal is having the power less than the power of the minimum detectable signal, then Radar cannot detect the target since it is beyond the maximum limit of the Radar's range.

Therefore, we can say that the range of the target is said to be maximum range when the received echo signal is having the power equal to that of minimum detectable signal. We will get the following equation, by substituting R=R_{MAX} and P_r =$P_{r(MIN)}$ in Equation (6).

$$R_{Max}= [\frac{P_t G S A_e}{(4\pi)^2 P_{r(MIN)}}]^{1/4} \qquad \text{Equation (7)}$$

Equation (7) represents the standard form of Radar range equation. By using the above equation, we can find the maximum range of the target.

Modified Forms of Radar Range Equation

We know the following relation between the Gain of directional Antenna, G and effective aperture, A_e.

$$G=4\pi A_e / \lambda^2 \qquad \text{Equation (8)}$$

Substitute, Equation (8) in Equation (7)

$$R_{Max}= [\frac{P_t S A_e}{(4\pi)^2 P_{r(MIN)}} (4\pi A_e / \lambda^2)]^{1/4}$$

$$\Rightarrow R_{Max}=[\frac{P_t S {A_e}^2}{4\pi\lambda^2 P_{r(MIN)}}]^{1/4} \qquad \text{Equation (9)}$$

Equation 9 represents the modified form of Radar range equation. By using the above equation, we can find the maximum range of the target.

We will get the following relation between effective aperture, Ae and the Gain of directional Antenna, G from Equation 8.

$$A_e=G\lambda^2 / 4\pi \qquad \text{Equation (10)}$$

Substitute, Equation 10 in Equation 7.

$$R_{Max}= [\frac{P_t G S}{(4\pi)^2 P_{r(MIN)}} (G\lambda^2 / 4\pi)]^{1/4}$$

$$\Rightarrow R_{Max}= [\frac{P_t G^2 S \lambda^2}{(4\pi)^3 P_{r(MIN)}}]^{1/4} \qquad \text{Equation (11)}$$

Equation (11) represents another modified form of Radar range equation. By using the above equation, we can find the maximum range of the target.

PERFORMANCE FACTOR OF RADAR-

The performance of a radar system can be judged by the following:

(1) The maximum range at which it can see a target of a specified size,

(2) The accuracy of its measurement of target location in range and angle,

(3) Its ability to distinguish one target from another,

(4) Its ability to detect the desired target echo when masked by large clutter echoes, unintentional interfering signals from other “friendly” transmitters, or intentional radiation from hostile jamming (if a military radar).

(5) Its ability to recognize the type of target,

(6) Its availability (ability to operate when needed), reliability, and maintainability.

Some of the major factors that affect performance are discussed below.

Transmitter power and antenna size-
The maximum range of a radar system depends in large part on the average power of its transmitter and the physical size of its antenna. (In technical terms, this is called the power-aperture product).

Receiver noise
The sensitivity of a radar receiver is determined by the unavoidable noise that appears at its input. At microwave radar frequencies, the noise that limits detectability is usually generated by the receiver itself (i.e., by the random motion of electrons at the input of the receiver) rather than by external noise that enters the receiver via the antenna. A radar engineer often employs a transistor amplifier as the first stage of the receiver even though lower noise can be obtained with more sophisticated (and more complex) devices.

Target size
The size of a target as “seen” by radar is not always related to the physical size of the object. The measure of the target size as observed by radar is called the radar cross section and is given in units of area (square meters). It is possible for two targets with the same physical cross-sectional area to differ considerably in radar size, or radar cross section.

Clutter
Echoes from land, sea, rain, snow, hill and birds are a noise to those who want to detect aircraft, ships, missiles, or other similar targets. Clutter echoes can seriously limit the capability of a radar system; thus, a significant part of radar design is devoted to minimizing the effects of clutter without reducing the echoes from desired targets. The Doppler frequency shift is the usual means by which moving targets are distinguished from the clutter of stationary objects.

Atmospheric effects

Rain and other forms of precipitation can cause echo signals that mask the desired target echoes. There are other atmospheric phenomena that can affect radar performance as well. The decrease in density of the Earth’s atmosphere with increasing altitude causes radar waves to bend as they propagate through the atmosphere.

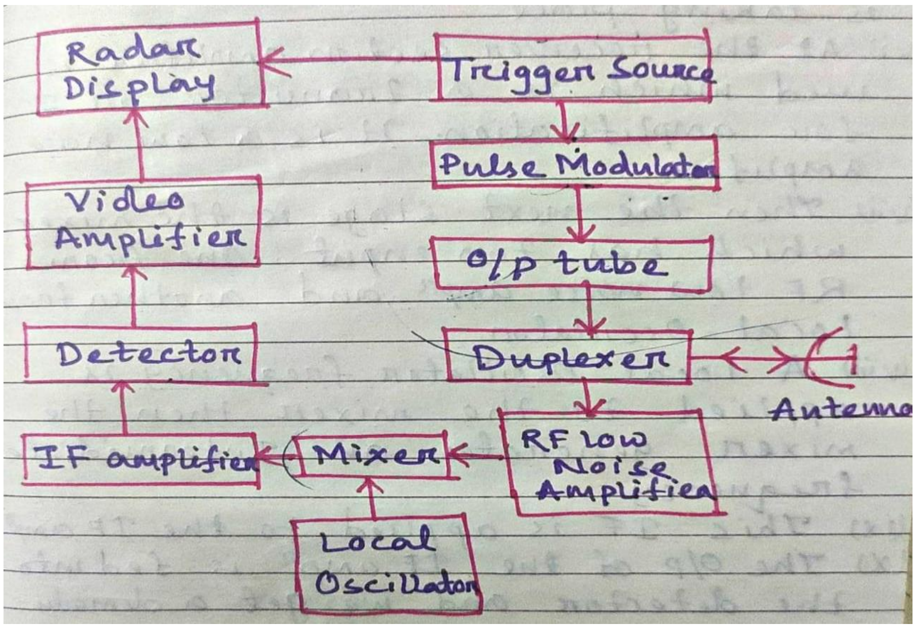
Interference
Signals from nearby radars and other transmitters can be strong enough to enter a radar receiver and produce spurious responses. Interference is not as easily ignored by automatic detection and tracking systems, however, and so some method is usually needed to

recognize and remove interference pulses before they enter the automatic detector and tracker of a radar.

WORKING PRINCIPLE OF PULSED RADAR SYSTEM-

PULSE RADAR SYSTEM –

Pulse Radar uses single Antenna for both transmitting and receiving of signals with the help of Duplexer. Following is the **block diagram** of Pulse Radar –



1) The trigger source provide pulse for the pulse modulator and radar display.
2) The modulator provides signal which is used to supply voltage for the o/p tube.
3) The o/p tube may be a magnetron oscillator or an amplifier such as the klystron or travelling wave tube depending on requirement.
4) Then the transmitter portion of the radar is connected with duplex which passes the o/p pulses to the antenna for transmission.
5) The receiver is also connected to the antenna through the duplexer. But it is connected when no transmission is taking place.
6) At the receiver section amplifier is used which is a transistor or IC for amplification. It is a low noise amplifier
7) A local oscillator frequency is applied to the mixer then the mixer generates an intermediate frequency.
8) This intermediate frequency is applied to the IF amplifier.
9) The output of the IF amplifier is fed into the detector and we get a demodulated signal.
10) The detector whose output is amplified by a video amplifier having the same bandwidth as the IF amplifier.
11) Its output then fed to a display unit to display the result.

MOVING TARGET INDICATION (MTI)-

Moving target indication (MTI) is a mode of operation of a radar to discriminate a target against the clutter. It describes a variety of techniques used to find moving objects, like an aircraft, and filter out unmoving ones, like hills or trees.

Early MTI systems generally used an acoustic delay line to store a single pulse of the received signal for exactly the time between broadcasts (the pulse repetition frequency). This stored pulse will be sent to the display along with the next received pulse. The result was that the signal from any objects that did not move mixed with the stored signal and became cancelled out. Only signals that changed, because they moved, remained on the display. These were subject to a wide variety of noise effects that made them useful only for strong signals, generally for aircraft or ship detection.

DOPPLER EFFECT-

- The **Doppler Effect** (or the **Doppler shift**) is the change in frequency of a wave in relation to an observer who is moving relative to the wave source
- A common example of Doppler shift is the change of pitch heard when a vehicle sounding a horn approaches and recedes from an observer. Compared to the emitted frequency, the received frequency is higher during the approach, identical at the instant of passing by, and lower during the recession.
- The reason for the Doppler Effect is that when the source of the waves is moving towards the observer, each successive wave crest is emitted from a position closer to the observer than the crest of the previous wave.
- Therefore, each wave takes slightly less time to reach the observer than the previous wave. Hence, the time between the arrivals of successive wave crests at the observer is reduced, causing an increase in the frequency. While they are traveling, the distance between successive wave fronts is reduced, so the waves "bunch together".
- Conversely, if the source of waves is moving away from the observer, each wave is emitted from a position farther from the observer than the previous wave, so the arrival time between successive waves is increased, reducing the frequency. The distance between successive wave fronts is then increased, so the waves "spread out".

C.W. RADAR-

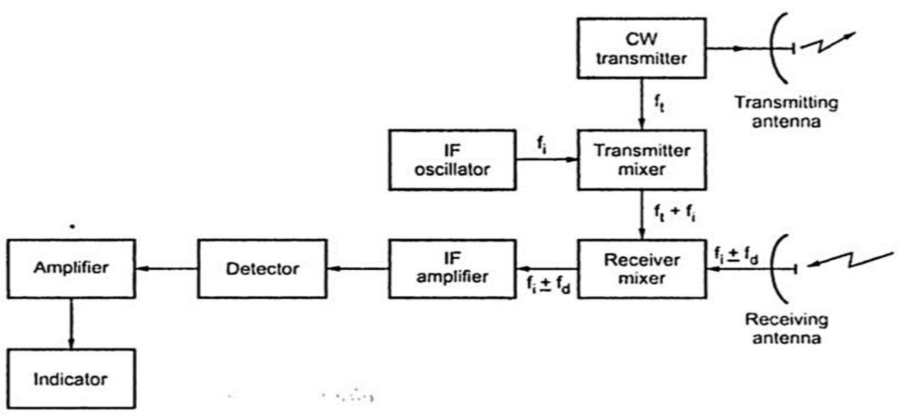


Fig. CW Doppler radar

- It is possible to detect moving targets by radiating unmodulated continuous wave (CW) energy instead of radiating in the form of pulses. Continuous wave (CW) radar makes use of the Doppler Effect for target speed measurements.
- The transmitter oscillator sends signal to the transmitting antenna i.e, f_t. The transmitting antenna transmit the signal to the space.
- A small portion of the transmitter output is mixed with output of the local oscillator and the sum is fed to the receiver mixer i.e., $f_t + f_i$.
- The receiver mixer also receives the Doppler shifted signal from its antenna i.e,$f_t \pm f_d$ and produces an output difference i.e, typically 30 MHz.
- The output of this mixer is amplified and demodulated. The signal from the detector is just the Doppler frequency.
- The detector output is fed into audio amplifier where it is amplified.
- Then the output of audio amplifier fed into the indicator to show the result.

Advantages of CW Doppler

1. CW Doppler radar has no blind speed.
2. CW Doppler radar is capable of giving accurate measurements of relative velocities.
3. CW Doppler radars are always on, they need low power and arc compact in size.
4. They can be used for small to large range with high degree of efficiency and accuracy.
5. The performance of radar is not affected by stationary object.

Disadvantages of CW Doppler Radar

1. The maximum range of CW Doppler radar is limited by the power that radar can radiate.
2. The target range cannot be calculated by CW Doppler radar.
3. There is possibility of ambiguous results when number of targets are more.

Applications of CW Doppler Radar

1. CW Doppler radars are used where only velocity information is of interest and actual range is not needed e.g. in police radar for catching cars travelling above the speed limit.
2. Measuring motion of waves on water level.
3. Traffic counters.
4. Intrusion alarm.
5. Runway monitors.
6. Cricket ball speed measurement.

AIDS TO NAVIGATION-

- Navigation is the art of guiding the movement of a craft from one point to another along a desired path. In older days long journeys over sea were accomplished with the knowledge of the movements of sun and various stars.
- These days most of the navigational work is done with electronic navigational aids. Electronic navigational aids are based on the use of EM waves to find the position of the craft.
- A no. of different types of navigational aids are available in the market. One of them is **LORAN.**
- **LORAN** is short form of Long Range Navigational Aid. It is based on the measurement of the difference in the time of arrival of EM waves from two transmitters to the receiver in the craft.
- Another one navigational aid is **radio range navigation.**

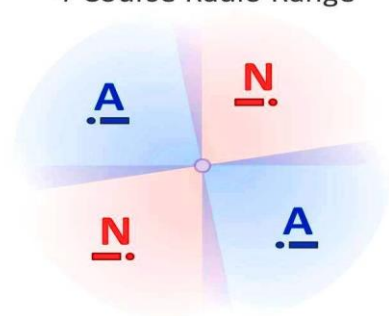
There are two types of radio range in operation-

(i) Low frequency or four course radio range
(ii) VHF Omni- directional radio range

Low frequency or four course radio range-

- The Low Frequency Radio range also known as the radio range was the main navigation system used by aircraft for instrument flying in the 1930s and 1940s until the use of VHF omnidirectional range.
- Based on a network of radio towers which transmitted directional radio signals the Low Frequency Radio signal defined a specific airways in the sky.
- Pilots navigated the Low Frequency Radio signal listening to a stream of automated "A" and "N" Morse codes.

4-Course Radio Range



VHF omnidirectional radio range-

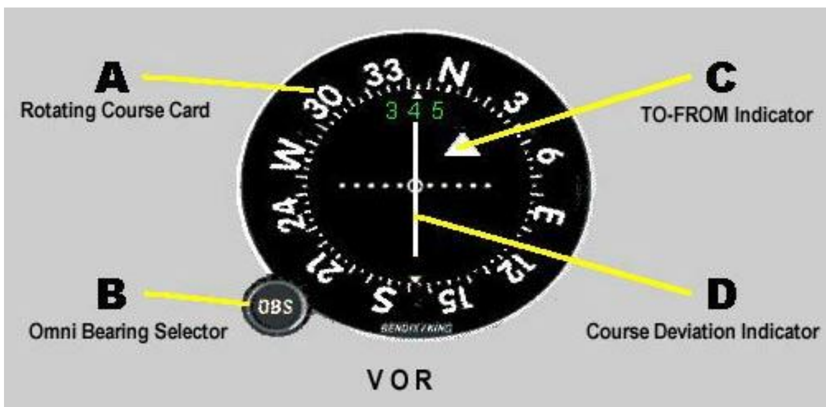
VOR (VHF Omni-Range) is the basic Electronic navigation that in use today. This VHF Omni-Range navigation method relies on the ground based transmitters which emitted signals to VOR receiver. The VOR system operates in the VHF frequency band, from 108.0 to 117.95 MHz. The reception of VHF signals is a line of sight situation. You must be on the minimum altitude of 1000 feet (AGL) above ground level in order to pick up an Omni signals service range.

<u>OPERATION</u>

The VOR facility at ground base transmits two signals at the same time. One signal is constant in all directions as a reference phase. Another signal, it is variable-phase signal and it rotates through 360 degrees, like the beam from the lighthouse. Both signals are in phase when the variable signal passes 360 degrees (reference to magnetic north) and they are 180 degrees out of phase when the rotating signal passes 180 degrees. The aircraft equipment receives both signals. The receiver will calculate the difference between the two signals, and interprets the result as a radial from the station to pilots on the aircraft.

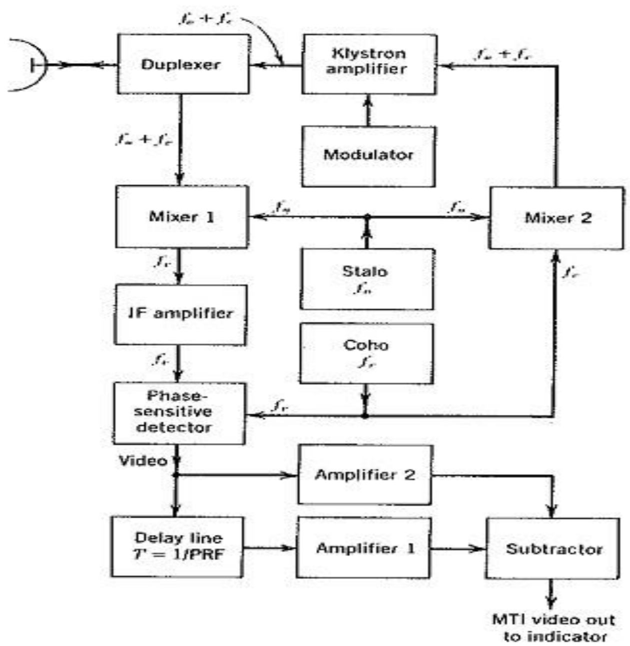
RADIALS: The two signals from VOR transmitter generate 360 lines like spokes in a wheel. Each line is called a Radial. VOR navigation equipment on the airplane will determine which of those 360 radials the airplane is on.

VOR INDICATOR



A : Rotating Course Card is calibrated from 0 to 360 degrees, which indicates the VOR bearing chosen as the reference to fly by pilot.
B : Omni Bearing Selector or OBS knob , used to manually rotate the course card to where the point to fly to.
C : TO-FROM indicator . The triangle arrow will point UP when flying to the VOR station. The arrow will point DOWN when flying away from the VOR station. A red flag replaces these TO-FROM arrows when the VOR is beyond reception range or the station is out.
D : Course Deviation Indicator (CDI). This needle moves left or right indicating the direction to turn the aircraft to return to course.
DOT : The horizontal dots at center are represent the aircraft away from the course . Each dot represent 2 degrees deviate from desired course.

MOVING TARGET INDICATION RADAR-



(1) The Moving Target Indicator Radar Block Diagram compares a set of received echoes with those received during the previous sweep. Those echoes whose phase has remained constant are then cancelled out. This applies to echoes due to stationary objects, but those due to moving targets do show a phase change; they are thus not cancelled nor is noise, for obvious reasons.

(2) The fact that clutter due to stationary targets is removed makes it much easier to determine which targets are moving and reduces the time taken by an operator to "take in" the display.

(3) It also allows the detection of moving targets whose echoes are hundreds of times smaller than those of nearby stationary targets and which would otherwise have been completely masked. MTI can be used with a radar using a power oscillator (magnetron) output.

(4) The transmitted frequency in the Moving Target Indicator Radar is the sum of the outputs of two oscillators, produced in mixer 2. The first is the stalo, or stable local oscillator (note that a good case can be made for using a varactor chain here).

(5) The second is the coho, or **coherent oscillator**, operating at the same frequency as the intermediate frequency and providing the **coherent signal**, which is used as will be explained. Mixers 1 and 2 are identical, and both use the same local oscillator (the stalo); thus phase relations existing in their inputs are preserved in their outputs.

(6) This makes it possible to use the Doppler shift at the JF, instead of the less convenient radio frequency $f_0 + f_c$.. The output of the IF amplifier and a reference signal from the coho are fed to the phase-sensitive detector, a circuit very similar to the phase discriminator.

(7) The coho is used for the generation of the RF signal, as well as for reference in the phase detector, and the mixers do not introduce differing phase shifts. The transmitted and reference signals are locked in phase and are said to be coherent; hence the name of the

coho. Since the output of this detector is phase-sensitive, an output will be obtained for all fixed or moving targets.

(8) The phase difference between the transmitted and received signals will be constant for fixed targets, whereas it will vary for moving targets. The phase shift is definitely not constant for Moving Target Indicator Radar.

(9) Each pulse that correspond to stationary targets are identical with each pulse, but those portions corresponding to moving targets keep changing in phase. It is thus possible to subtract the output for each pulse from the preceding one, by delaying the earlier output by a time equal to the pulse interval, or 1/PRF.

(10) Since the delay line also attenuates heavily and since signals must be of the same amplitude if permanent echoes are to cancel, an amplifier follows the delay line. To ensure that this does not introduce a spurious phase shift, an amplifier is placed in the undelayed line, which has exactly the same response characteristics (but a much lower gain) than amplifier 1. The delayed and undelayed signals are compared in the subtractor and displayed.

AIR CRAFT LANDING SYSTEM-

An *instrument landing system* operates as a ground-based instrument approach system that provides precision lateral and vertical guidance to an aircraft approaching and landing on a runway, using a combination of radio signals and, in many cases, high-intensity lighting arrays to enable a safe landing during instrument meteorological conditions (IMC), such as low ceilings or reduced visibility due to fog, rain, or blowing snow.

An instrument approach procedure chart is published for each ILS approach to provide the information needed to fly an ILS approach during instrument flight rules (IFR) operations. A chart includes the radio frequencies used by the ILS components and the prescribed minimum visibility requirements.

An aircraft approaching a runway is guided by the ILS receivers in the aircraft by performing modulation depth comparisons. Many aircraft can route signals into the autopilot to fly the approach automatically. An ILS consists of two independent sub-systems. The localizer provides lateral guidance; the glide slope provides vertical guidance.

Localizer

The localizer station for runway

A localizer an antenna array normally located beyond the departure end of the runway and generally consists of several pairs of directional antennas.

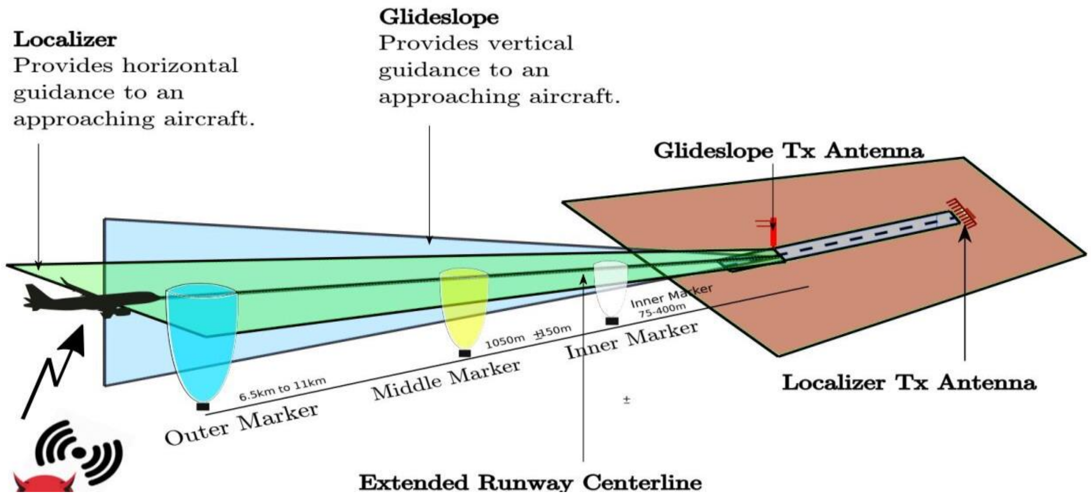
The localizer will allow the aircraft to turn and match the aircraft with the runway. After that, the pilots will activate approach phase.

Glide slope of ILS (G/S)

Glide slope station for runway

The pilot has to correct to the left and a little upwards.

The pilot controls the aircraft so that the glide slope indicator remains centered on the display to ensure the aircraft is following the glide path of approximately 3° above horizontal (ground level) to remain above obstructions and reach the runway at the proper touchdown point (i.e., it provides vertical guidance).



Limitations

Due to the complexity of ILS localizer and glide slope systems, there are some limitations. Localizer systems are sensitive to obstructions in the signal broadcast area, such as large buildings or hangars. Glide slope systems are also limited by the terrain in front of the glide slope antennas. If terrain is sloping or uneven, reflections can create an uneven glide path, causing unwanted needle deflections. Additionally, since the ILS signals are pointed in one direction by the positioning of the arrays, glide slope supports only straight-line approaches with a constant angle of descent. Installation of an ILS can be costly because of siting criteria and the complexity of the antenna system.

Variant

- Instrument guidance system (IGS) is a modified ILS to accommodate a non-straight approach.

<u>NAVSAT-</u>

- NAVSAT system is a system that uses satellites to provide autonomous geo spatial positioning.
- It allows small electronic receivers to determine their location to high precision using time signals transmitted along a line of sight by radio from satellites.
- The system can be used for providing position, navigation or for tracking the position of something fitted with a receiver.
- The signals also allow the electronic receiver to calculate the current local time to high precision which allows time synchronization.
- A navigational satellite system with global coverage may be termed a global navigation satellite system (GNSS).
- The first satellite navigation system was Transit, a system deployed by the US military in the 1960s.
- Transit's operation was based on the Doppler Effect. The satellites travelled on well-known paths and broadcast their signals on a well-known radio frequency.

- The received frequency will differ slightly from the broadcast frequency because of the movement of the satellite with respect to the receiver.
- By monitoring this frequency shift over short time interval, the receiver can determine its location.
- Modern systems are more direct. The satellite broadcasts a signal that contains orbital data.

<u>GLOBAL POSITIONING SYSTEM-</u>

- The GPS project was started by the U.S. Department of Defense in 1973, with the first prototype spacecraft launched in 1978 and the full constellation of 24 satellites operational in 1993. Originally limited to use by the United States military, civilian use was allowed from the 1980s.
- The GPS is a satellite based navigation system made up of at least 24 satellites.
- GPS works in any weather condition, anywhere in the world, 24 hours a day with no subscription charges.

Working-

- GPS satellites circle the earth twice a day in a precise orbit. Each satellite transmits an unique signal and orbital parameters that allow GPS devices to decode and compute the precise location of the satellite. GPS receiver use this information to calculate a user's exact location.
- The GPS receiver measures the distance to satellite by the amount of time it takes to receive a transmitted signal.
- With distance measurements from a few more satellites the receiver can determine a user's position and display it electronically to measure the running route and find a way to destination.
- To calculate the 2D position (latitude & longitude) and track movement, a GPS receiver must be locked onto the signal of at least 3 satellites.
- With 4 or more satellites, the receiver can determine the 3D position i.e. latitude, longitude and altitude.
- Once our position has been determined, the GPS can calculate other information. The 31 satellites that currently makeup the GPS system are orbiting the earth about 1200miles above us.
- This satellites are constantly moving, making 2 complete orbits in less than 24 hours. They travel at speeds of roughly 7000miles an hours.

SIMPLE RADAR PROBLEMS-

<u>Problem: - 1</u>

Calculate the maximum range of RADAR for the following specifications:-

- Peak power transmitted by the RADAR, P_t= 250KW
- Gain of transmitting Antenna , G=4000
- Effective aperture of the receiving Antenna, A_e =4m^2
- Radar cross section of the target, S= 25m^2
- Power of minimum detectable signal, $P_{r_{(MIN)}}$ = 10^{-12}W

Solution :-

We can use the following standard form of RADAR range equation in order to calculate the maximum range of Radar for given specifications.

$$R_{Max}= \left[\frac{P_t G S A_e}{(4\pi)^2 P_{r(MIN)}}\right]^{1/4}$$

Substitute all the given parameters in above equation.

$$R_{Max}=\left[\frac{(250\ X\ 10^3)(4000)(25)(4)}{(4\pi)^2(10^{-12})}\right]^{1/4}$$

=> R_{Max} =158 KM

Therefore, the maximum range of Radar for given specification is 158KM.

<u>Problem 2</u>

Calculate the maximum range of Radar for the following specifications.

- Operating frequency, f=10GHZ
- Peak power transmitted by the Radar, P_t= 400KW
- Effective aperture of the receiving Antenna, $A_e = 5m^2$
- Radar cross section of the target, S=30m^2
- Power of minimum detectable signal, $P_{r(MIN)} = 10^{-10}W$

Solution

We know the following formula for operating wavelength, λ in terms of operating frequency, f.

λ = C/f

Substitute, C=3 X 10^8m/sec and f=10GHZ in above equation.

$$\lambda = \frac{3\ X\ 10^8}{10\ X\ 10^9}$$

λ =0.03m

So, the operating wavelength, λ is equal to 0.03m, when the operating frequency, f is 10GHZ

We can use the following modified form of Radar range equation in order to calculate the maximum range of Radar for given specifications.

$$R_{Max}= \left[\frac{P_t S A_e{}^2}{(4\pi\lambda^2 P_{r(MIN)}}\right]^{1/4}$$

Substitute, the given parameters in the above equation.

$$R_{Max}=\left[\frac{(400X\ 10^3)(30)(5^2)}{4\pi(0.003)^2(10)^{-10}}\right]^{1/4}$$

$$=> R_{Max} = 128\ KM$$

UNIT-2: SATELLITE COMMUNICATION

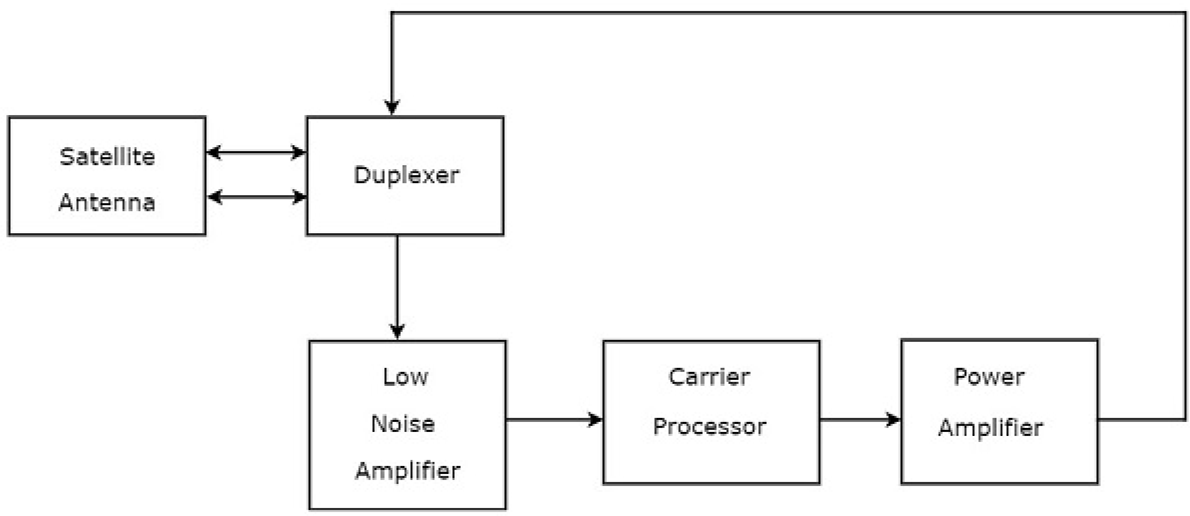
Satellite Transponder-

Transponder performs the functions of both transmitter and receiver (Responder) in a satellite. Hence, the word 'Transponder' is obtained by the combining few letters of two words, Transmitter **(Trans)** and Responder **(ponder)**.

Block diagram of Transponder

Transponder performs mainly **two functions**. Those are amplifying the received input signal and translates the frequency of it. In general, different frequency values are chosen for both uplink and down link in order to avoid the interference between the transmitted and received signals.

The **block diagram** of transponder is shown in below figure.



We can easily understand the operation of Transponder from the block diagram itself. The function of each block is mentioned below.

- **Duplexer** is a two-way microwave gate. It receives uplink signal from the satellite antenna and transmits downlink signal to the satellite antenna.
- **Low Noise Amplifier** (LNA) amplifies the weak received signal.
- **Carrier Processor** performs the frequency down conversion of received signal (uplink). This block determines the type of transponder.
- **Power Amplifier** amplifies the power of frequency down converted signal (down link) to the required level.

Types of Transponders

Basically, there are **two types** of transponders. Those are Bent pipe transponders and Regenerative transponders.

Bent Pipe Transponders

Bent pipe transponder receives microwave frequency signal. It converts the frequency of input signal to RF frequency and then amplifies it.

Bent pipe transponder is also called as repeater and **conventional transponder**. It is suitable for both analog and digital signals.

Regenerative Transponders

Regenerative transponder performs the functions of Bent pipe transponder. i.e., frequency translation and amplification. In addition to these two functions, Regenerative transponder also performs the demodulation of RF carrier to baseband, regeneration of signals and modulation.

Regenerative transponder is also called as processing transponder. It is suitable only for digital signals. The main **advantages** of Regenerative transponders are improvement in Signal to Noise Ratio (SNR) and have more flexibility in implementation.

KEPLER'S LAW-

Kepler's laws of planetary motion can be stated as follows:

Kepler First law – The Law of Orbits

According to Kepler's first law, all the planets revolve around the sun in elliptical orbits having the sun at one of the foci. The point at which the planet is close to the sun is known as perihelion and the point at which the planet is farther from the sun is known as aphelion.

It is the characteristics of an ellipse that the sum of the distances of any planet from two foci is constant. The elliptical orbit of a planet is responsible for the occurrence of seasons.

Kepler's Second Law – The Law of Equal Areas

As the orbit is not circular, the planet's kinetic energy is not constant in its path. It has more kinetic energy near perihelion and less kinetic energy near aphelion implies more speed at perihelion and less speed (V_{min}) at aphelion.

" The radius vector from the sun to a planet sweeps equal areas in equal times."

Kepler's Third Law – The Law of Periods

According to Kepler's law of periods, the square of the time period of revolution of a planet around the sun in an elliptical orbit is directly proportional to the cube of its semi-major axis.

$$T^2 \propto a^3$$

GEO, MEO AND LEO-

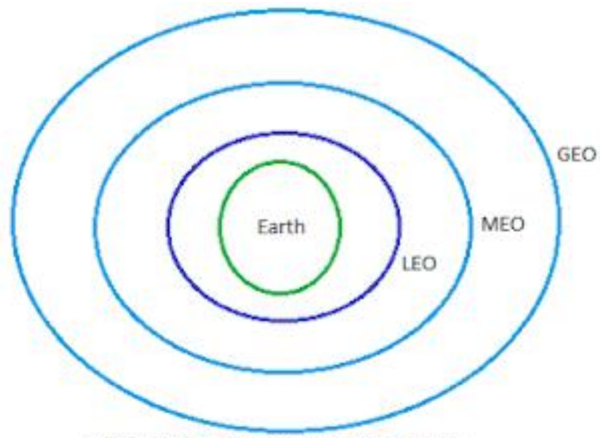


Fig: Satellite orbits according to altitude

Geo-Stationary Earth Orbit (GEO)-

- These satellites have almost a distance of 35,786 km to the earth. This orbit has the special characteristic that the apparent position of the satellite in the sky when viewed by a ground observer does not change, the satellite appears to stand still in the sky. This is because the satellite's orbital period is same as the rotation rate of the earth.
- E.g. All radio and TV broadcast satellites etc, are launched in this orbit.
- The signal delay approximately 480ms.

Advantages of Geo-Stationary Earth Orbit

1. It is possible to cover almost all parts of the earth with just 3 geo satellites.
2. Antennas need not be adjusted every now and then but can be fixed permanently.
3. The life-time of a GEO satellite is quite high usually around 15 years.

Disadvantages of Geo-Stationary Earth Orbit

1. Larger antennas are required for northern/southern regions of the earth.
2. High buildings in a city limit the transmission quality.
3. High transmission power is required.
4. These satellites cannot be used for small mobile phones.
5. Fixing a satellite at Geo stationary orbit is very expensive.

Medium Earth Orbit (MEO)-

- The orbital path spreads over 2000km to 35,786km above earth in circular path is called medium earth orbit and the satellite moves around this orbit is called MEO satellite.
- MEO satellites are visible for much longer periods of time than LEO satellites usually between 2-8 hrs.
- It is useful for GPS.
- The signal delay approximately 100ms.
- These orbits have moderate number of satellites.

Advantages of Medium Earth Orbit

1. Compared to LEO system, MEO requires only a dozen satellites.

2. Simple in design.

3. Requires very few handovers.

4. MEO satellites have a larger coverage area than LEO satellite.

Disadvantages of Medium Earth Orbit

1. Satellites require higher transmission power.

2. Special antennas are required.

3. MEO satellites don't maintain a stationary distance from the earth.

Low Earth Orbit (LEO)-

- LEO satellites operate at a distance of about 160-2000 km. Because of their low altitude, these satellites are only visible for few time.
- It is useful for military observation, metrology, atmospheric studies and ground studies.
- The signal delay approximately 20ms.

Advantages of Low Earth Orbit

1. The antennas can have low transmission power of about 1 watt.

2. The delay of packets is relatively low.

3. Useful for smaller foot prints.

Disadvantages of Low Earth Orbit

1. If global coverage is required, it requires at least 50-200 satellites in this orbit.

2. Special handover mechanisms are required.

3. These satellites involve complex design.

4. Very short life: Time of 5-8 years. Assuming 48 satellites with a life-time of 8 years each, a new satellite is needed every 2 months.

5. Data packets should be routed from satellite to satellite.

CALCULATION OF HEIGHT OF A GEO STATIONARY SATELLITE-

The gravitational force between the satellite and the Earth is in the radial direction and its magnitude is given by the Newton's equation

$F = GMm/r^2$ (1)

Where G is the gravitational constant, M and m are the masses of the Earth and the satellite respectively and r is the radius of the orbit.

In case of the circular motion the net force equals mass times acceleration, where acceleration can be calculated by $\omega^2 r$, where ω is the angular rate of rotation also known as angular velocity.

Thereby,

$F = ma = m\omega^2 r$. (2)

The angular velocity is given by

$\omega = 2\pi/T$, (3)

where T is the period for one rotation.

We substitute (3) into the equation (2) and we get

$F = m4\pi^2 r/T^2$. (4)

Now we can use the equations (4) and (1) to find the following formula

$m4\pi^2 r/T^2 = GMm/r^2$

or

$r^3 = GMT^2/4\pi^2$. (5)

We substitute the values and we get

$r^3 = 6.67*10^{-11}*5.972*10^{24}*86400^2/4\pi^2$

and finally,

$r = 4.22*10^7$ m.

The radius of the Earth is $6.37*10^6$ m.

We can calculate the height h above the Earth's surface by subtracting the radius of the Earth from the radius of the orbit.

$h = 4.22*10^7 – 6.37*10^6 = 3.583*10^7$ m.

CALCULATION OF VELOCITY OF A GEO STATIONARY SATELLITE-

A geostationary orbit is a circular orbit directly above the Earth's equator approximately 35,786 km above ground. Any point on the equator plane revolves about the Earth in the same direction and with the same period as the Earth's rotation.
The period of the satellite is one day or approximately 24 hours. To find the speed of the satellite in orbit we use Newton's law of gravity and his second law of motion along with that we know about centripetal acceleration. The inward and outward forces on the satellite must equal each another (by Newton's first law of motion).
$F_{centripetal} = F_{centrifugal}$
By Newton's second law of motion:
F = ma

$m_S . a_g = m_S . a_c$

Where:
m_s – Mass of satellite
a_g – Gravitational acceleration
a_c – Centrifugal acceleration

The centripetal acceleration provided by Earth's gravity:

$a_g = (M_e . G) / r^2$

Where:
M_e - Mass of Earth in kilograms (5.9742×10^{24} kg)
G - Gravitational constant (6.6742 x 10-11 N m2 kg-2 = 6.6742×10^{-11} m^3 s^{-2} kg^{-1})

Magnitudes of the centrifugal acceleration derived from orbital motion:
$a_c = \omega^2 . r$

Where;
ω - Angular velocity in radians per second.
r - Orbital radius in meters as measured from the Earth's center of mass.

$a_c = (v/r)^2$.r

$$a_c = \frac{v^2}{r}$$

From the relationship

$$m_{S\,.}\ (M_e.\,\mathrm{G})\ /\ r^2 = m_S\,.\frac{v^2}{r}$$

$F_{centripetal} = F_{centrifugal}$
we note that the mass of the satellite, m_S appears on both sides, geostationary orbit is independent of the mass of the satellite.

$$(M_e.\,\mathrm{G})\ /\ r^2 = \frac{v^2}{r}$$

$$\mathrm{V} = \sqrt{(M_e.\,\mathrm{G})\ /\ r}$$

r (Orbital radius) = Earth's equatorial radius + Height of the satellite above the Earth surface
r = 6,378 km + 35,780 km
r = 42,158 km
r = 4.2158 x 107 m

$$\mathrm{V} = \sqrt{\frac{(5.9742 \times 10^{24}\mathrm{kg} \times 6.6742 \times 10^{-11}\ \mathrm{N}m^2\ Kg^{-2})}{4.2158 \times 10^7 m}}$$

$$\mathrm{V} = 3.0754 \times 10^3\ \mathrm{m}s^{-1}$$

Speed of the satellite is 3.0754 x 103 m/s

<u>CALCULATION OF ROUND TRIP TIME DELAY OF GEOSTATIONARY SATELLITE-</u>

Most communications satellites are located in the Geostationary Orbit (GSO) at an altitude of approximately 35,786 km above the equator. At this height the satellites go around the earth in a west to east direction at the same angular speed at the earth's rotation, so they appear to be almost fixed in the sky to an observer on the ground.

The time for one satellite orbit and the time for the earth to rotate is 1 sidereal day or 23 h 56 m 4 seconds.

Radio waves go at the speed of light which is 300,000 km per second.

If you are located on the equator and are communicating with a satellite directly overhead then the total distance, single hop (up and down) is nearly 72,000 km so the time delay is **240 ms**. mS means millisecond or 1 thousandth of a second so 240 mS is just under a quarter of a second.

A geostationary satellite is visible from a little less than one third of the earth's surface and if you are located at the edge of this area the satellite appears to be just above the horizon. The distance to the satellite is greater and for earth stations at the extreme edge of the coverage area, the distance to the satellite is approx 41756 km. If you were to communicate with another similarly located site, the total distance is nearly 84,000 km so the end to end delay is almost **280 mS**, which is a little over quarter of a second.

Advantages

- At 35,786km is a special orbit for Geo-stationary satellites. Since they orbit the earth in the same time that it takes the earth to spin once on its axis, it means that the satellite can stay in roughly the same spot over the earth.
- The main advantage is that they appear to stay in the same spot in the sky, meaning that satellite dishes can be fixed onto them, rather than having to track them accross the sky.
- Their orbit is quite high, so they do not experience as much atmospheric drag - their orbits may be a lot more stable.

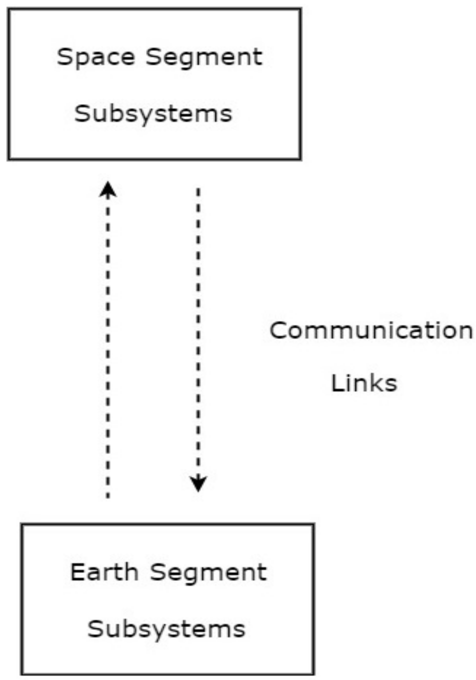
Limitations

- Geostationary satellites have two major limitations. First, because the orbital zone is an extremely narrow ring in the plane of the equator, the number of satellites that can be maintained in geostationary orbits without mutual conflict (or even collision) is limited. Second, the distance that an electromagnetic (EM) signal must travel to and from a geostationary satellite is a minimum of 71,600 kilometers or 44,600 miles. Thus, a latency of at least 240 milliseconds is introduced when an EM signal, traveling at 300,000 kilometers per second (186,000 miles per second), makes a round trip from the surface to the satellite and back.
- There are two other, less serious, problems with geostationary satellites. First, the exact position of a geostationary satellite, relative to the surface, varies slightly over the course of each 24-hour period because of gravitational interaction among the satellite, the earth, the sun, the moon, and the non-terrestrial planets. As observed from the surface, the satellite wanders within a rectangular region in the sky called the box. The box is small, but it limits the sharpness of the directional pattern, and therefore the power gain, that earth-based antennas can be designed to have. Second, there is a dramatic increase in background EM noise when the satellite comes near the sun as observed from a receiving station on the surface, because the sun is a powerful source of EM energy. This effect, known as solar fade, is a problem only within a few days of the equinoxes in late March and late September. Even then, episodes last for only a few minutes and take place only once a day.

SATELLITE SUB SYSTEM-

In satellite communication system, various operations take place. Among which, the main operations are orbit controlling, altitude of satellite, monitoring and controlling of other subsystems.

A satellite communication consists of mainly two **segments**. Those are space segment and earth segment. So, accordingly there will be two types of subsystems namely, space segment subsystems and earth segment subsystems. The following **figure** illustrates this concept.



As shown in the figure, the **communication** takes place between space segment subsystems and earth segment subsystems through communication links.

Space Segment Subsystems

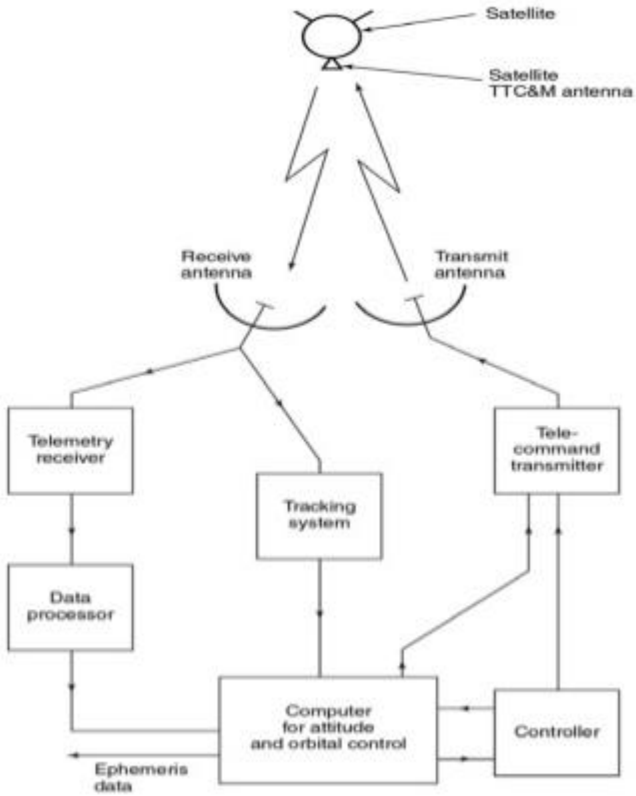
The subsystems present in space segment are called as space segment subsystems. Following are the **space segment subsystems**.

- AOC Subsystem
- TTCM Subsystem
- Power and Antenna Subsystems
- Transponders

Earth Segment Subsystems

Earth segment is also called as ground segment.

Earth segment performs mainly two functions. Those are transmission of a signal to the satellite and reception of signal from the satellite. **Earth stations** are the major subsystems that are present in earth segment.



<u>AOC subsystem-</u>

Altitude and Orbit Control **(AOC)** subsystem consists of rocket motors, which are capable of placing the satellite into the right orbit, whenever it is deviated from the respective orbit. We can make this AOC subsystem into the following **two parts**.

- Altitude Control Subsystem
- Orbit Control Subsystem

Altitude Control Subsystem

Altitude control subsystem takes care of the orientation of satellite in its respective orbit.

Orbit Control Subsystem

Orbit control subsystem is useful in order to bring the satellite into its correct orbit, whenever the satellite gets deviated from its orbit.

The TTCM subsystem present at earth station monitors the position of satellite. If there is any change in satellite orbit, then it sends a signal regarding the correction to Orbit control subsystem. Then, it will resolve that issue by bringing the satellite into the correct orbit.

In this way, the **AOC subsystem** takes care of the satellite position in the right orbit and at right altitude during entire life span of the satellite in space.

TTCM-

- Telemetry, Tracking, Commanding and Monitoring **(TTCM)** subsystem is present in both satellite and earth station. In general, satellite gets data through sensors. So, Telemetry subsystem present in the satellite sends this data to earth station(s). Therefore, TTCM subsystem is very much necessary for any communication satellite in order to operate it successfully.
- It is the responsibility of satellite operator in order to control the satellite in its life time, after placing it in the proper orbit. This can be done with the help of **TTCM subsystem**.

TTCM subsystem consists of **three parts**.

- Telemetry and Monitoring Subsystem
- Tracking Subsystem
- Commanding Subsystem

Telemetry and Monitoring Subsystem

- The word **'Telemetry'** means measurement at a distance. Mainly, the following operations take place in 'Telemetry'.
  - Generation of an electrical signal, which is proportional to the quantity to be measured.
  - Encoding the electrical signal.
  - Transmitting this code to a far distance.
- **Telemetry subsystem** present in the satellite performs mainly two functions –
  - receiving data from sensors, and
  - Transmitting that data to an earth station.

Satellites have quite a few sensors to monitor different parameters such as pressure, temperature, status and etc., of various subsystems.

Telemetry subsystem is a remote controlled system. It sends monitoring data from satellite to earth station.

Tracking Subsystem

Tracking subsystem is useful to know the position of the satellite and its current orbit. Satellite Control Center **(SCC)** monitors the working and status of space segment subsystems with the help of telemetry downlink. And, it controls those subsystems using command uplink.

We know that the **tracking subsystem** is also present in an earth station. It mainly focusses on range and look angles of satellite.

The **tracking subsystem** that is present in an earth station keeps tracking of satellite, when it is released from last stage of Launch vehicle. It performs the functions like, locating of satellite in initial orbit and transfer orbit.

Commanding Subsystem

- Commanding subsystem is necessary in order to launch the satellite in an orbit and its working in that orbit. This subsystem adjusts the altitude and orbit of satellite, whenever there is a deviation in those values.
- It also controls the communication subsystem. This **commanding subsystem** is responsible for turning ON / OFF of other subsystems present in the satellite based on the data getting from telemetry and tracking subsystems.

Power Systems

Power system is a vital subsystem, which provides the power required for working of a satellite. Mainly, the solar cells (or panels) and rechargeable batteries are used in these systems.

Solar Cells

The **solar cells** produce electrical power (current) from incident sunlight. Therefore, solar cells are used primarily in order to provide power to other subsystems of satellite.

Individual solar cells generate very less power. So, in order to generate more power, group of cells that are present in an array form can be used.

Solar Arrays

There are two **types of solar arrays** that are used in satellites. Those are cylindrical solar arrays and rectangular solar arrays or solar sail.

- **Cylindrical solar arrays** are used in spinning satellites. Only part of the cylindrical array will be covered under sunshine at any given time. Due to this, electric power gets generated from the partial solar array. This is the drawback of this type.
- The drawback of cylindrical solar arrays is overcome with **Solar sail**. This one produce more power because all solar cells of solar sail are exposed to sun light.

Rechargeable Batteries

During eclipses time, it is difficult to get the power from sun light. So, in that situation the other subsystems get the power from **rechargeable batteries**. These batteries produce power to other subsystems during launching of satellite also.

These batteries charge due to excess current, which is generated by solar cells in the presence of sun light.

Antenna Subsystems

Satellite antennas perform **two types** of functions. Those are receiving of signals, which are coming from earth station and transmitting signals to one or more earth stations based on the requirement. In other words, the satellite antennas receive uplink signals and transmit downlink signals.

SATELLITE FREQUENCY ALLOCATION & FREQUENCY BANDS-

Frequency allocation (or **spectrum allocation** or spectrum management) is the allocation and regulation of the electromagnetic spectrum into radio frequency bands, which is normally done by governments in most countries. Because radio propagation does not stop at national boundaries, governments have sought to harmonies the allocation of RF bands and their standardization.

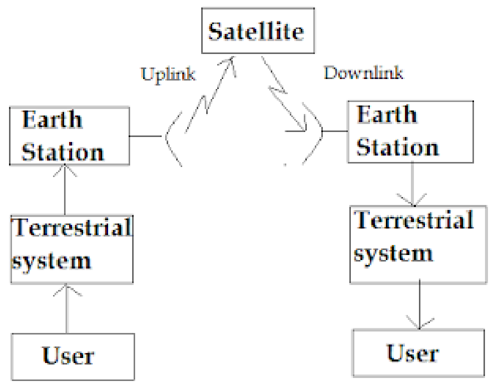
C band **Satellite** will have about 500 MHz of **bandwidth** between upper and lower frequency limits. This 500 MHz bandwidth is divided into 12 channels of 36 MHz bandwidth. 4MHz of guard bands incorporated in between this channels to minimize adjacent channel interference.

Satellite applications include FSS, BSS and MSS. FSS stands for Fixed Service Satellite, BSS stands for Broadcast Service Satellite and MSS stands for Mobile Service Satellite.

Frequency Band	Frequency Range(GHz)	Bandwidth(GHz)	Applications
L band	1-2	1	MSS
S band	2-4	2	MSS,NASA
C band	4-8	4	FSS
X band	8-12.5	4.5	meteorological satellite and FSS military
Ku band	12.5-18	5.5	FSS,BSS
K band	18-26.5	8.5	FSS,BSS

Ka band	26.5-40	13.5	FSS

GENERAL STRUCTURE OF SATELLITE LINK SYSTEM-



- This consists of a satellite in space that links many earth stations on the ground. The user is connected the earth station through terrestrial network. This terrestrial network may be a telephone switch or a dedicated link to the earth station.
- The user generates the base-band signal that is processed and transmitted to the satellite at the earth station.
- Thus, the satellite may be thought of as a large repeater in space that receives the modulated RF carriers in its up-link (earth to space) frequency spectrum from all the earth stations in the network, amplifies these carriers and re-transmits them back to the earth in the down-link (space to earth) frequency spectrum which is different from the up-link frequency spectrum in order to avoid the interference.
- The signal at the receiving earth station is processed to get back the baseband signal which is then sent to the user through a terrestrial network.
- On the guidelines of WARC-1979, commercial communication satellite use a frequency band of 500MHz bandwidth near 6GHz for up-link transmission and another 500MHz bandwidth near 4GHz for down-link transmission (i.e. 6/4 GHz band). In fact an up-link of 5.725 to 7.075GHz and a down-link of 3.4 to 4.8GHZ is used.
- The 500MHz allocation is usually divided into 12 channels of approximately 40MHz each and the transmit power per channel is typically of the order of 5 to 10W. This allows each of up to 12 transponders to carry one TV channel or 1500 analog FM voice circuits.

- This 6/4 band have been the most popular because they offer the fewest propagation problems and also RF components for these bands have been readily available.
- Rain attenuation is also not much serious at these bands. Sky noise is also low at 4GHz and so it is possible to build receiving system with low noise at 4GHz.
- With the overcrowding of GEO satellites at 6/4 GHz band, 14/12 GHz band is also being used in commercial communication satellites.

A third band where extremely high capacities are potentially available is the 30/20 GHz band.

Transponder-

A **transponder** receives and transmits radio signals at a prescribed frequency range. After receiving the signal a **transponder** will at the same time broadcast the signal at a different frequency. ... **Transponders** are **used in** satellite communications and in location identification and navigation systems.

Uplink and downlink-

The communication going from a satellite to ground is called downlink, and when it is going from ground to a satellite it is called uplink. When an uplink is being received by the spacecraft at the same time a downlink is being received by Earth, the communication is called two-way. If there is only an uplink happening, this communication is called upload. If there is only a downlink happening, the communication is called one-way.

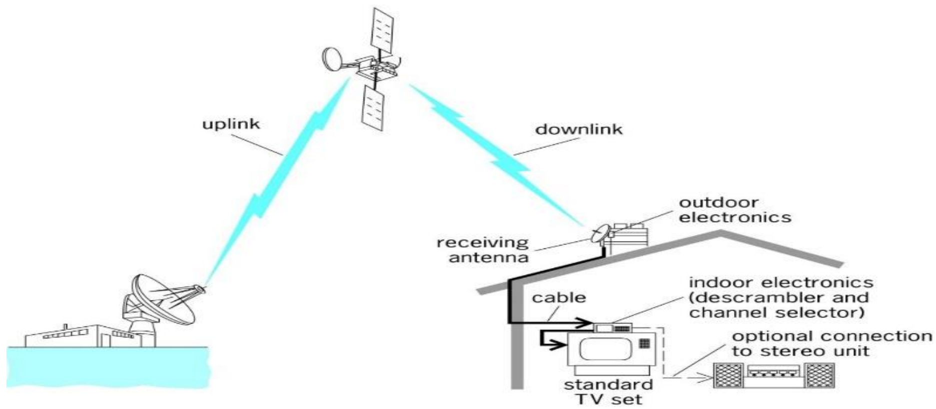
Crosslink-

A crosslink is a communications link between two satellites, typically serving a relay function in a constellation. For example, in the following illustration, there is a crosslink between Satellite1 and Satellite2, serving to relay communications between two ground stations. The crosslink is needed because the orbits of the satellites in the constellation are too low for a single satellite to have simultaneous access to both ground stations.

<u>DIRECT BROADCASTING SATELLITE SYSTEMS-</u>

- Systems for transmitting television and other program material via satellite directly to ind ividual homes and businesses. Direct broadcasting satellite (DBS) systems operate at micr owave frequencies, in a portion of the Ku band; in North and South America these systems operate in the frequency range 12.2–12.7 GHz.
- DBS systems use a satellite in geostationary orbit to receive television signals sent up fro m the Earth's surface, amplify them, and transmit them back down to the surface.
- The satellite also shifts the signal frequency, so that a signal sent up to the satellite in the 1 7.3-17.8GHz uplink band is transmitted back down in the 12.212.7GHz downlink band.
- The downlink signal is picked up by a receive antenna located atop an individual home or office; these antennas are usually in the form of a parabolic dish, but flat square Phased array antennas are sometimes used, and may eventually become commonplace.

- The receive antenna may be permanently pointed at the satellite, which is at a fixed point in the sky, in a geostationary orbit.



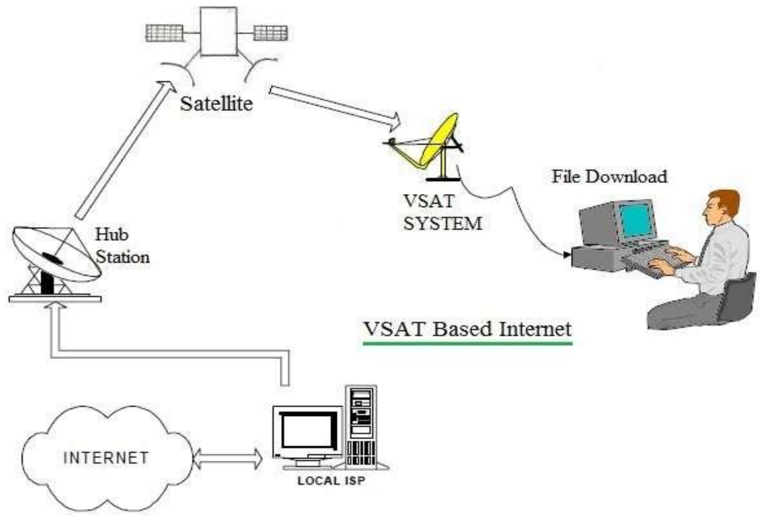
Direct broadcasting satellite system

Operation-

- A direct broadcast satellite system, a satellite broadcast company encodes TV signals that it receives from its partners and transmits them to an orbiting satellite.
- Although a satellite broadcast company may have nos. of satellites in orbit, they may either transmit data directly from multiple transmitters to each satellite or simply use one satellite to share information with another.
- Once the appropriately positioned satellite obtains the data that it needs it retransmits the information to all residences in a specific area. If a residence in that area has paid for services with the satellite broadcast company it will be equipped with a satellite dish and a receiver box that is able to receive the signals and decode them, allowing all users in the residence to access TV services.

VSAT-

- Low cost business terminals with small antennas (generally less than 2 meters in diameter) are often termed Very Small Aperture Terminals (VSATs).



- These are usually perceived as being two way data terminals, though strictly speaking many of the systems used for data broadcast are really one-way VSATs. Taking the USA as an example, approximately half of all installed VSATs are only used for one way data links.
- ETSI take a different definition for a VSAT as a one or two-way terminal used in a star, mesh or point to point network. Antenna size is restricted to being less than or equal to 3.8 m at Ku band and 7.8 m at C band.
- The majority of VSAT antennas range from 75cm to 1.2cm.
- VSATs access satellites in geosynchronous orbit or geostationary orbit to relay data from small remote earth stations to other terminals or master earth station “hubs”.
- A VSAT end user needs a box that interfaces between the user’s computer and an outside antenna i.e. transceiver.
- The transceiver receives or sends a signal to a satellite transponder in the sky.
- The satellite sends and receives signals from an earth station computer that acts as a hub for the system.
- Each end user is interconnected with the hub station via the satellite in a star topology.
- For one end user to communicate with another, each transmission has to first go to the hub station which retransmits it via the satellite to the other end user’s VSAT.
- VSAT handles data, voice and video signals

Advantages of VSAT

1. Installation: VSAT services are deployed in hours or minutes.
2. Coverage: It can be available anywhere with clear line of sight between VSAT antenna disc and satellite over the earth. It is popular in hilly areas where other mode of communication is either not available or difficult to install.
3. Price: VSAT terminals are cheaper.

4. Upgradation: It is flexible to add a VSAT site and increase the bandwidth as per future requirements.
5. Service charges: It depends on the bandwidth allocated as per user requirements.
6. VSAT provides same quality of service and speed at all the locations across the entire VSAT network.
7. VSAT services are independent of other wired and wireless mediums used as transmission network service provider. Hence it is a great backup system which is available during disaster and emergency situations.
8. VSAT terminals and indoor/outdoor hardware can be installed on truck or van and can be used even in mobility conditions.

Disadvantages of VSAT

1. As mentioned it requires clear Line of Sight between VSAT dish and satellite in the space.
2. The malfunctioning of satellite and Hub station (in case of star topology) will lead to disruption of VSAT services. To avoid this situation, redundant systems and switch over units are needed to have backup systems available for hot switching in faulty situations. But this increases overall cost of the VSAT system as a whole.
3. Latency for packet transmission from source to destination is higher due to distance of satellite from earth is about 36000 Km. Latency further increases in star topology of VSAT, as it requires two hops to reach a final destination.
4. VSAT services get affected in bad weather conditions.
5. As information transmitted by VSAT goes over the air till it reaches destination, it is prone to intrusion by hackers. Hence encryption-decryption units are needed to have secure communication. This increases the overall VSAT terminal cost.

MULTIPLE ACCESSING & ITS TYPES-

Multiple access is a **technique** that lets **multiple** mobile users share the allotted spectrum in the most effective manner. In computer networks and telecommunications, the **multiple access** method permits various terminals to connect to the same **multi**-point transmission medium to transmit over it and share its capacity.

There are several different ways to allow access to the channel. These includes mainly the following –

- Frequency division multiple-access (FDMA)
- Time division multiple-access (TDMA)
- Code division multiple-access (CDMA)
- Space division multiple access (SDMA)
- Spread Spectrum Multiple Access (SSMA)

Frequency Division Multiple Access (FDMA)

FDMA is the basic technology for advanced mobile phone services. The features of FDMA are as follows.

- FDMA allots a different sub-band of frequency to each different user to access the network.
- If FDMA is not in use, the channel is left idle instead of allotting to the other users.
- FDMA is implemented in Narrowband systems and it is less complex than TDMA.
- Tight filtering is done here to reduce adjacent channel interference.
- The base station BS and mobile station MS, transmit and receive simultaneously and continuously in FDMA.

Time Division Multiple Access (TDMA)

In the cases where continuous transmission is not required, there TDMA is used instead of FDMA. The features of TDMA include the following.

- TDMA shares a single carrier frequency with several users where each users makes use of non-overlapping time slots.
- Data transmission in TDMA is not continuous, but occurs in bursts. Hence handsoff process is simpler.
- TDMA uses different time slots for transmission and reception thus duplexers are not required.
- TDMA has an advantage that is possible to allocate different numbers of time slots per frame to different users.
- Bandwidth can be supplied on demand to different users by concatenating or reassigning time slot based on priority.

Code Division Multiple Access (CDMA)

Code division multiple access technique is an example of multiple access where several transmitters use a single channel to send information simultaneously. Its features are as follows.

- In CDMA every user uses the full available spectrum instead of getting allotted by separate frequency.
- CDMA is much recommended for voice and data communications.
- While multiple codes occupy the same channel in CDMA, the users having same code can communicate with each other.
- CDMA offers more air-space capacity than TDMA.
- The hands-off between base stations is very well handled by CDMA.

Space Division Multiple Access (SDMA)

Space division multiple access or spatial division multiple access is a technique which is MIMO (multiple-input multiple-output) architecture and used mostly in wireless and satellite communication. It has the following features.

- All users can communicate at the same time using the same channel.
- SDMA is completely free from interference.
- A single satellite can communicate with more satellites receivers of the same frequency.
- The directional spot-beam antennas are used and hence the base station in SDMA, can track a moving user.
- Controls the radiated energy for each user in space.

Spread Spectrum Multiple Access (SSMA)

Spread spectrum multiple access (SSMA) uses signals which have a transmission bandwidth whose magnitude is greater than the minimum required RF bandwidth.

There are two main types of spread spectrum multiple access techniques −

- Frequency hopped spread spectrum (FHSS)
- Direct sequence spread spectrum (DSSS)

Frequency Hopped Spread Spectrum (FHSS)

This is a digital multiple access system in which the carrier frequencies of the individual users are varied in a pseudo random fashion within a wideband channel. The digital data is broken into uniform sized bursts which is then transmitted on different carrier frequencies.

Direct Sequence Spread Spectrum (DSSS)

This is the most commonly used technology for CDMA. In DS-SS, the message signal is multiplied by a Pseudo Random Noise Code. Each user is given his own code word which is orthogonal to the codes of other users and in order to detect the user, the receiver must know the code word used by the transmitter.

The combinational sequences called as **hybrid** are also used as another type of spread spectrum. **Time hopping** is also another type which is rarely mentioned.

Since many users can share the same spread spectrum bandwidth without interfering with one another, spread spectrum systems become **bandwidth efficient** in a multiple user environment.

<u>TIME DIVISION MULTIPLE ACCESS (TDMA)-</u>

TDMA system divide the bandwidth into time slots and in each slot only one user is allowed to either transmit or receive. TDMA systems transmit data in a buffer- and- burst method, thus the transmission for any user in non-continuous.

In a TDMA frame the preamble contains the address and synchronization information that both the base station and the subscribers used to identify each other. Guard times are utilized to allow synchronization of the receivers between different slots and frames. The features of TDMA include the following-

1. TDMA shares a single carrier frequency with several users, where each user makes use of non-overlapping time slots.
2. Data transmission for users of a TDMA system is not continuous, but occurs in bursts. This results in low battery consumption, since the subscriber transmitter can be turned off when not in use.
3. Because of discontinuous transmissions in TDMA the hand off process is much simpler for a subscriber, since it is able to listen for other base stations during ideal time slots.
4. TDMA uses different time slots for transmission and reception thus duplexers are not required.
5. Adaptive equalizer is usually necessary in TDMA systems, since the transmission rates are very high.
6. In TDMA the guard time should be minimized.
7. High synchronization overhead is required in TDMA systems because of burst transmissions. TDMA transmission are slotted and this requires the receiver to be synchronized for each data burst. Guard slots are necessary to separate users and this results in a TDMA system having large overheads.
8. TDMA has an advantage is that it is possible to allocate different numbers of time slots per frame to different users. Thus bandwidth can be supplied on demand to different users.

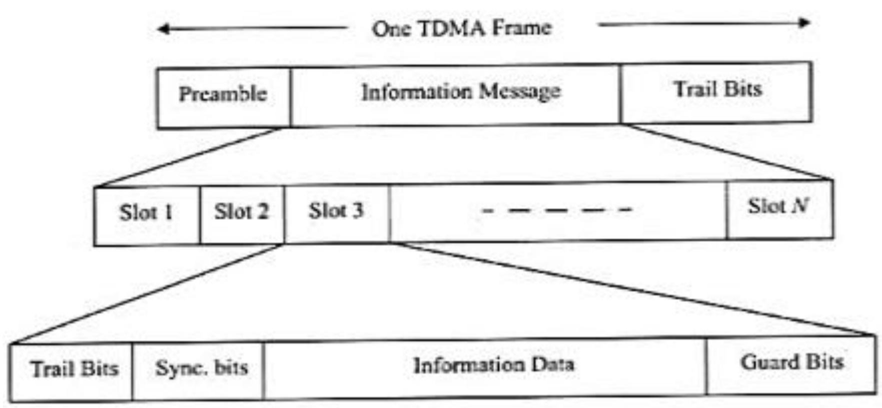


Fig: TDMA frame structure

Advantages of TDMA:

- TDMA can easily adapt to the transmission of data as well as voice communication.
- It has the ability to carry 64 kbps to 120 Mbps of data rates.

- No interference from simultaneous transmission.
- TDMA is the cost-effective technology to convert an analogue system to digital.
- Share a single carrier frequency with multiple users
- TDMA provides the user with extended battery life since transmitting the only portion of the time during conversations
- Flexible bit rate
- TDMA separates users according to time ensures that there will be no interference from the simultaneous transmission.
- TDMA allows the operator to do services like fax, voice, data, SMS as well as applications such as multimedia and video conferencing.

<u>Disadvantages of TDMA:</u>

- In TDMA each user has a predefined time slot so that users roaming from one cell to another are not allotted a time slot. Thus, if all the time slots in the next cell are already occupied, a cell might well be disconnected. In the same way, if all the time slots in the cell in which a user happens to be in are already occupied, a user will not receive a dial tone.
- High synchronization overhead.
- Frequency/slot allocation is to be complex in TDMA.
- Equalization was necessary for high data rates.

<u>CODE DIVISION MULTIPLE ACCESS (CDMA)-</u>

It is cellular technology in which there are two main systems i.e, Base Station (BS) and users. In CDMA systems the narrow band message signal is multiplied by a very large bandwidth signal called the spreading signal. The spreading signal is a pseudo noise or pseudo random code. All users in a CDMA system use the same carrier frequency and may transmit simultaneously. Each user has its own pseudo random code. The receiver performs operation to detect only the specific desired code word. All other code appears as noise. For detection of the message signal, the receiver need to know the code word used by the transmitter.

Advantages-

The use of CDMA offers several advantages that's why CDMA technology is adopted by many of the 3G cellular telecommunications systems. CDMA technology in mobile communication possesses so many advantages. The following advantages are a few of them:

1. **Security**: It is difficult for hackers to tap the CDMA signals. Hence it is more secure.
2. **Improvement in capacity and security**: In CDMA technology data and voice packets are separated using codes, and then transmitted by using a wide range of frequencies. Because more space is allocated for data in CDMA, this standard has become attractive for 3G high-speed mobile internet use.
3. **Improvement in hand over/ hand off:** By using CDMA technology, it is easy for a terminal to communicate with two base stations at once. In case of this, old link is to be broken when the new

one is firmly established. This provides improvement in terms of the reliability of hand over/hand off from one base station to another.

CDMA technology has been used in 3G telecommunication systems in one form or the other. CDMA has become successful in every aspect, and it has enabled improvements need to be gained over the previous technologies used in 2G systems.

Disadvantages-

- In CDMA, orthogonal codes are used by mobile subscriber. Orthogonality between the codes need to be maintained in order to recover the data. The subscribers which are farthest from base station will have more attenuation and hence will lose the orthogonality and hence it will be difficult to recover data.
- CDMA uses soft handoff. In this type of handoff mobile needs to establish connection with the new target cell before disconnecting itself from serving cell. This procedure is more complex compare to the hard handoff type.
- Increase in number of users will decrease the overall quality of service.
- Near far problem is causes in CDMA system. This requires close control of transmit powers of CDMA handsets.
- Self-jamming is observed in CDMA system due to loss of orthogonality of PN codes.

Applications of CDMA technology

- Due to inherent advantages of CDMA over TDMA and FDMA such as user capacity, soft hand offs and security, etc., CDMA emerges as a winner in the battle of wireless technology and services. CDMA allows far greater development and the use of broad band devices such as wireless laptop modems, GPS system units and other innovative devices.
- For business purpose, CDMA supports in providing high speed push to talk and push to email services. Push to talk gives mobile an ability to be used as a walky-talky device. These services are exempted from the service charges imposed by the operators making CDMA cost effective.
- CDMA is considered as the highest mode of wireless communications, and is responsible for gibing fast and safe mode of data exchange such as 3G. Recently, CDMA has merged with the GSM technology to give a high-speed 4G or LTE internet services.

SATELLITE APPLICATION-

COMMUNICATION SATELLITE-

- A communication satellite is an artificial satellite that amplifies radio telecommunication signal using a transponder, it creates a communication channel between a transmitter and receiver at different locations on earth.

- Communication satellites are used for TV, radio, telephone, internet and military applications. There are over 2000 communication satellites in earth orbit used by both private and government organization.
- Wireless communication uses EM waves to carry signals. These waves requires line of sight propagation and are thus obstructed by the curvature of the earth. The purpose of communication satellite is to relay the signal around the curve of the earth allowing communication between widely separated points.
- Communication satellites use a wide range of radio range of radio and microwave frequencies. To avoid signal interference international organizations have regulations for which frequency ranges or bands are allocated to certain organization. This allocation of bands minimizes the risk of signal interference.
- **Advantages-**
  1. Flexibility- Satellites are able to provide communications in a variety of ways.
  2. Mobility- Satellite communications are able to reach all areas of the globe depend upon the type of satellite system in use.
- **Disadvantages-**
  1. Cost- Satellites are very expensive to build, place in orbit and then maintain.
  2. Propagation delay- As distances are very much greater than those involved with terrestrial systems, propagation delay can be an issue, especially for satellites using geostationary orbits.
- **Applications-**
  1. Telecommunication- Satellite systems have been able to provide data communication links over large distances.
  2. Satellite phones- The concept of using mobile phone from anywhere on the globe is one that has many applications. There are still many areas where coverage is not available. In these situations satellite phones are of great use.
  3. Direct Broadcast- Direct broadcast satellite is a type of satellite application that is used to broadcast data directly to a residence or commercial office.

<u>MSAT-</u>

- On April 20, 1996, the Canadian firm TMI Communications launched MSat, a powerful new communications satellite. MSat was Canada's first satellite designed to serve mobile users, especially those in remote areas out of the reach of conventional communication systems.
- Most telephone calls are made over phone lines that physically connect the user to a station that transmits the signal to its destination.
- This is only possible if the users are always near their telephones. Cellular phones, which serve mobile users, only operate when the phone is within range of a microwave tower. Because these signals, however, must travel in a straight line, they cannot travel around the curvature of the Earth or past large obstacles like mountains.
- The MSat system is revolutionary because any user equipped with a communicator the size of a briefcase can transmit directly to a satellite.

- This mini-terminal will allow users to make voice phone calls, send email or faxes, and obtain accurate information about their position. Any vessel, vehicle, aircraft, or remote operation can easily be equipped with one.
- One of the reasons the new satellite telephone service works for remote locations is that the new satellite broadcasts a signal of 600 watts, eight times stronger than any previous commercial satellite.
- This means that the signal no longer requires a large dish-like antenna to pick up the signal. Instead, the MSat communicator can receive and transmit the digital signals.
- There are many uses for this new service. For example, fire fighters trying to put out a forest fire in a remote location can communicate with airplanes bringing water to drop on the fire.
- Air ambulances can communicate with medical experts on the ground. Truck drivers can communicate with their head office while they're travelling.
- MSat is paired with the American satellite AMSC. Together, the two satellites provide communications capability to almost every square inch of North and Central America, and the Caribbean. Also, the two satellites serve as backup for one another.
- The two satellites were built through joint contracts with the Canadian firm Spar Aerospace Ltd. and the American firm Hughes Communications Inc.

DIGITAL SATELLITE RADIO-

- Satellite Radio is defined by the International Telecommunication Union as a broadcasting satellite service.
- The satellite's signals are broadcast nationwide, across a much wider geographical area than terrestrial radio stations and the service is primarily intended for the motor vehicles.
- It is available by subscription, mostly commercial free and offers subscribers more stations and a wider variety of programming options than terrestrial radio.
- Ground stations transmit signals to the satellites which are 35,786 Km. above the equator.
- The satellites send the signals back down to radio receivers in cars and homes.
- This signal contains scrambled broadcasts, along with Meta data about each specific broadcast.
- The signals are unscrambled by the radio receiver modules, which display the broadcast information.
- Satellite radio uses the 2.3GHz S band in North America for nationwide digital radio broadcasting.
- In other parts of the world, satellite radio uses part of the 1.4GHz L band.
- Satellite radio subscribers purchase a receiver and pay a monthly subscription fee to listen to programming.
- They can listen through built in or portable receivers in automobiles, in the home and office with a portable receiver equipped to connect the receivers to a stereo system.
- There has been three major satellite radio companies: World space, Sirius Satellite radio and XM satellite radio.

GPS TRANSMITTER AND RECEIVER

The current GPS consists of three major segments. These are the space segment, a control segment, and a user segment. The U.S. Space Force develops, maintains, and operates the space and control segments. GPS satellites broadcast signals from space, and each GPS receiver uses these signals to calculate its three-dimensional location (latitude, longitude, and altitude) and the current time.

Space segment-

The space segment (SS) is composed of 24 to 32 satellites, or Space Vehicles (SV), in medium Earth orbit, and also includes the payload adapters to the boosters required to launch them into orbit. The GPS design originally called for 24 SVs, eight each in three approximately circular orbits, but this was modified to six orbital planes with four satellites each.

The space segment consists of 24 satellites revolving the earth at 1200 miles in altitude. This high altitude allows the signals to cover a greater area. This satellites are arranged in their orbits. So a GPS receiver on earth an always receive a signal from at least 4 satellites at any given time.

Control Segment-

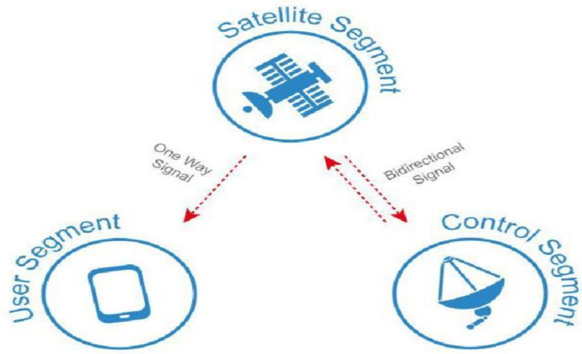
The control segment (CS) is composed of:

1. a master control station (MCS),
2. an alternative master control station,
3. four dedicated ground antennas, and
4. Six dedicated monitor stations.

The control segment tracks the satellites and then provides them with orbital and time information. The control segment consists of 4 un-manned control station and one master control station. The 4 un-manned station receive data from the satellites and then transmit that information to the master control station, where it is corrected and sent back to GPS satellites.

User Segment-

- The user segment consists of the user and their GPS receiver and the no. of simultaneous user is limitless.
- When GPS receiver is turned ON, it first downloads the orbit information of all satellites, this process for the first time can take some time, but once this information is downloaded, it is stored in the receiver memory for future use.
- The GPS receiver can calculate the distance of moving body, by multiplying the velocity of the transmitted signal by the time.
- To determine time part, the receiver matches the satellite's transmitted code to its own code and by comparing them determines how much it needs to delay its code to match the satellite's code. This delay time is multiplied by velocity to get the distance.



APPLICATION-

Aviation

Almost all modern aircraft are fitted with multiple GPS receivers. This provides pilots (and sometimes passengers) with a real-time aircraft position and map of each flight's progress. GPS also allows airline operators to pre-select the safest, fastest and most fuel-efficient routes to each destination.

Marine When high accuracy GPS is fitted to boats and ships, it allows captains to navigate through unfamiliar harbours, shipping channels and waterways.

Farming

Farmers rely on repeat planting season after season to maximize their crop productions. By putting GPS receivers on tractors and other agricultural equipment, farmers can map their plantations and ensure that they return to exactly the same areas when sewing their seeds in future. This strategy also allows farmers to continue working in low visibility conditions such as fog and darkness.

2 Science

Scientists use GPS technology to conduct a wide range of experiments and research, ranging from biology to physics to earth sciences.

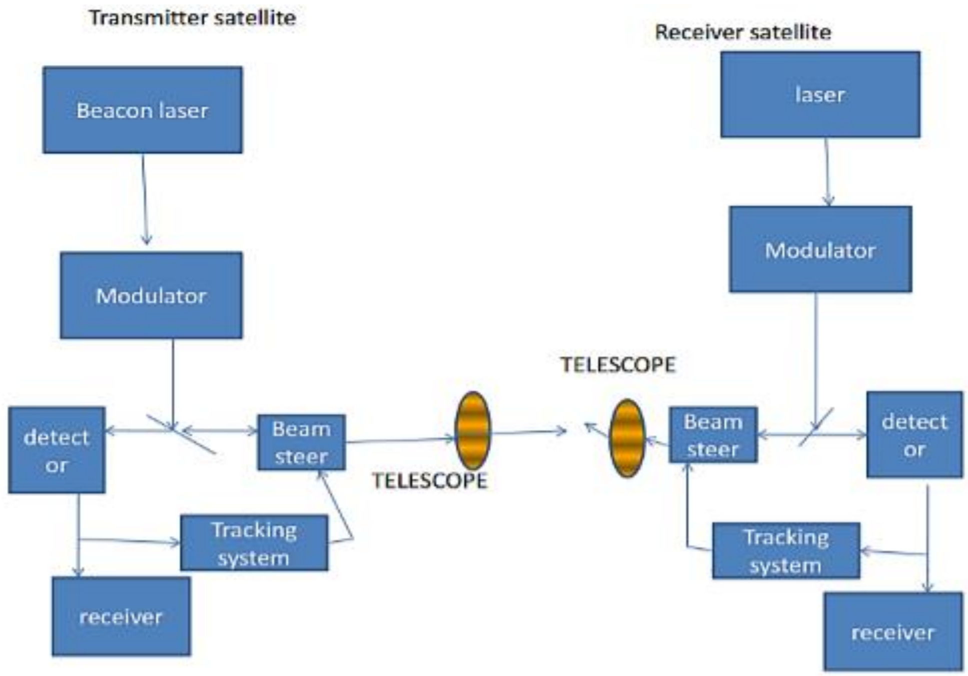
Surveying

Surveyors are responsible for mapping and measuring features on the earth's surface and under water with high accuracy. GPS can either be setup over a single point to establish a reference marker, or it can be used in a moving configuration to map out the boundaries of various features. This data can then be transferred into mapping software to create very quick and detailed maps for customers.

Military

The GPS system was originally developed by the United States Department of Defence for use by the US military, but was later made available for public use. GPS is used to map the location of vehicles and other assets on various battlefields in real time, which helps to manage resources and protect soldiers on the ground. GPS technology is also fitted to military vehicles and other hardware such as missiles, providing them with tracking and guidance to various targets at all times of the day and in all weather conditions.

OPTICAL SATELLITE LINK TRANSMITTER & RECEIVER-



Optical Link Model

In comparison with radio links, optical links have specific characteristics. Two aspects should be indicated:

The small diameter of the telescope is typically of the order 0:3 m. In this way, one is freed from congestion problems and aperture blocking of other antennas in the payload.

The narrowness of the optical beam is typically 5 micro radians. Notice that this width is several orders of magnitude less than that of a radio beam and this is an advantage for protection against interference between systems. But it is also a disadvantage since the beam width is much less than the precision of satellite attitude control (typically 0.1_ or 1.75 mrad). Consequently an advanced pointing device is necessary; this is probably the most difficult technical problem.

There are three basic phases to optical communications:

Acquisition:

The beam must be as wide as possible in order to reduce the acquisition time. But this requires a high-power laser transmitter. A laser of lower mean power can be used which emits pulses of high peak power with a low duty cycle. The beam scans the region of space where the receiver is expected to be located. When the receiver receives the signal, it enters a tracking phase and transmits in the direction of the received signal. On receiving the return signal from the receiver, the transmitter also enters the tracking phase. The typical duration of this phase is 10 seconds.

Tracking:

The beams are reduced to their nominal width. Laser transmission becomes continuous. In this phase, which extends throughout the following, the pointing error control device must allow for movements of the platform and relative movements of the two satellites. In addition, since the relative velocity of the two satellites is not zero, a lead-ahead angle exists between the receiver line of sight and the transmitter line of sight. As demonstrated below, the lead-ahead angle is larger than the beam width and must be accurately determined.

Communications:

Information is exchanged between the two ends.

UNIT-3: OPTICAL FIBER COMMUNICATION

Principle of optical fiber communication:

Optical communication is any type of communication in which light is used to carry the signal to the remote end, instead of electrical current. Optical communication relies on optical fibers to carry signals to their destinations. A modulator/demodulator, a transmitter/receiver, a light signal and a transparent channel are the building blocks of the optical communications system.

At the transmitting end, the transmitted information (such as voice) is first converted into an electrical signal, and then modulated onto the laser beam emitted by the laser, so that the intensity of the light changes with the amplitude (frequency) of the electrical signal, and passes through the optical fiber. The principle of total reflection is transmitted; at the receiving end, after receiving the optical signal, the detector converts it into an electrical signal, and after demodulation, restores the original information.

Optical communication utilizes the principle of total reflection. When the injection angle of light satisfies certain conditions, light can form total reflection in the optical fiber, thereby achieving the purpose of long-distance transmission. The light guiding properties of an optical fiber are based on total reflection of the light ray at the core and cladding interfaces, limiting light transmission in the core. There are two types of light in the fiber, namely meridional and oblique rays. The meridional rays are the light rays on the meridional plane, and the oblique rays are the light that is transmitted through the fiber axis.

ADVANTAGE AND DISADVANTAGE OF OPTICAL FIBER OVER METALLIC CABLES-

Basic for comparison	Optical Fiber	Metallic Cable
Basic of transmission	Transmission of signal is in optical form.	Transmission of signal is in electrical form.
Composition of the cable	Glass and plastic	Plastic, metal foil and metal wire.
Losses in cable	Dispersion bending, absorption and attenuation	Resistive, radiated and dielectric loss
Efficiency	High	Low

Cost	Highly expensive	Less expensive
Bending effect	Can affect the signal transmission	Bending of wire does not affect the signal transmission
Data transmission rate	2Gbps	44.736 Mbps
Installation of the cable	Difficult	Easy
Bandwidth provided	Very high	Moderately high
External magnetic field	Does not affect the cable	Affects the cable
Noise immunity	High	Intermediate
Diameter of the cable	Smaller	Larger
Weight of the cable	Lighter	Heavier

ELECTROMAGNETIC FREQUENCY-

Electromagnetic radiation phenomena with wavelengths ranging from as long as one meter to as short as one millimeter are called microwaves; with frequencies between 300 MHz (0.3 GHz) and 300 GHz.

WAVE LINE SPECTRUM-

Spectrum, in optics, the arrangement according to wavelength of visible, ultraviolet, and infrared light. An instrument designed for visual observation of spectra is called a spectroscope; an instrument that photographs or maps spectra is a spectrograph. Spectra may be classified according to the nature of their origin, *i.e.,* emission or absorption. An emission spectrum consists of all the radiations emitted by atoms or molecules, whereas in an absorption spectrum, portions of a continuous spectrum (light containing all wavelengths) are missing because they have been absorbed by the medium through which the light has passed; the missing wavelengths appear as dark lines or gaps.

The spectrum of incandescent solids is said to be continuous because all wavelengths are present. The spectrum of incandescent gases, on the other hand, is called a line spectrum because only a few wavelengths are emitted. These wavelengths appear to be a series of parallel lines because a slit is used as the light-imaging device. Line spectra are characteristic of the elements that emit the radiation. Line spectra are also called atomic spectra because the lines represent wavelengths radiated from atoms when electrons change from one energy level to another. Band spectra is the name given to groups of lines so closely spaced that each group appears to be a band, *e.g.,* nitrogen spectrum. Band spectra, or molecular spectra, are produced by molecules radiating their rotational or vibrational energies, or both simultaneously.

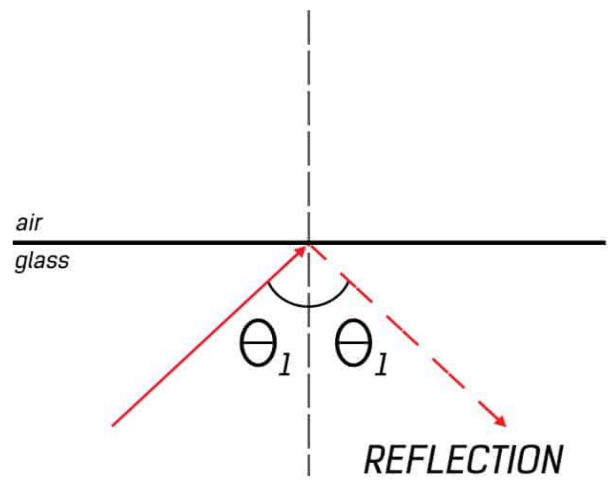
PRINCIPLES OF LIGHT TRANSMISSION IN A FIBER USING RAY THEORY-

The basic laws of ray theory are quite self-explanatory

- In a homogeneous medium, light rays are straight lines.

- Light may be absorbed or reflected
- Reflected ray lies in the plane of incidence and angle of incidence will be equal to the angle of reflection.
- At the boundary between two media of different refractive indices, the refracted ray will lie in the plane of incidence. Snell's Law will give the relationship between the angles of incidence and refraction.

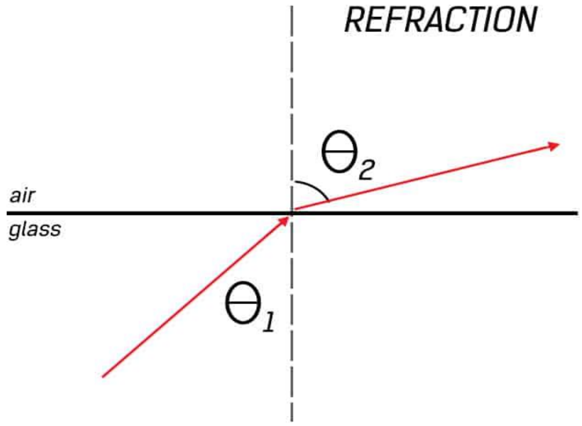
REFLECTION OF LIGHT



The Law of Reflection states that the angle of incidence must be equal to the angle of reflection.

Reflection depends on the type of surface on which light is incident. An essential condition for reflection to occur with glossy surfaces is that the angle made by the incident ray of light with the normal at the point of contact should be equal to the angle of reflection with that normal. The *images* produced from this reflection have different properties according to the shape of the surface. For example, for a flat mirror, the image produced is upright, has the same size as that of the object, and is equally distanced from the surface of the mirror as the real object. However, the properties of a parabolic mirror are different, and so on. Reflection is also possible in one other scenario, even if the surface is not reflective.

REFRACTION OF LIGHT-



Refraction is the bending of light in a particular medium due to the speed of light in that medium. The speed of light in any medium can be given by

$$v = \frac{c}{n}$$

Here n is the **refractive index** of that medium. When a ray of light is incident at the interface of two media with different refractive indices, it will bend either towards or away from the normal depending on the refractive indices of the media. According to **Snell's law**, refraction can be represented as

$$n_1 \sin(\theta_1) = n_2 \sin(\theta_2)$$

n_1 = refractive index of first medium

θ_1 = angle of incidence

n_2 = refractive index of second medium

θ_2 = angle of refraction

For $n_1 > n_2$, θ_2 is always greater than θ_1. **Or** to put it in different words, light moving from a medium of high refractive index (glass) to a medium of lower refractive index (air) will move away from the normal.

Total internal reflection

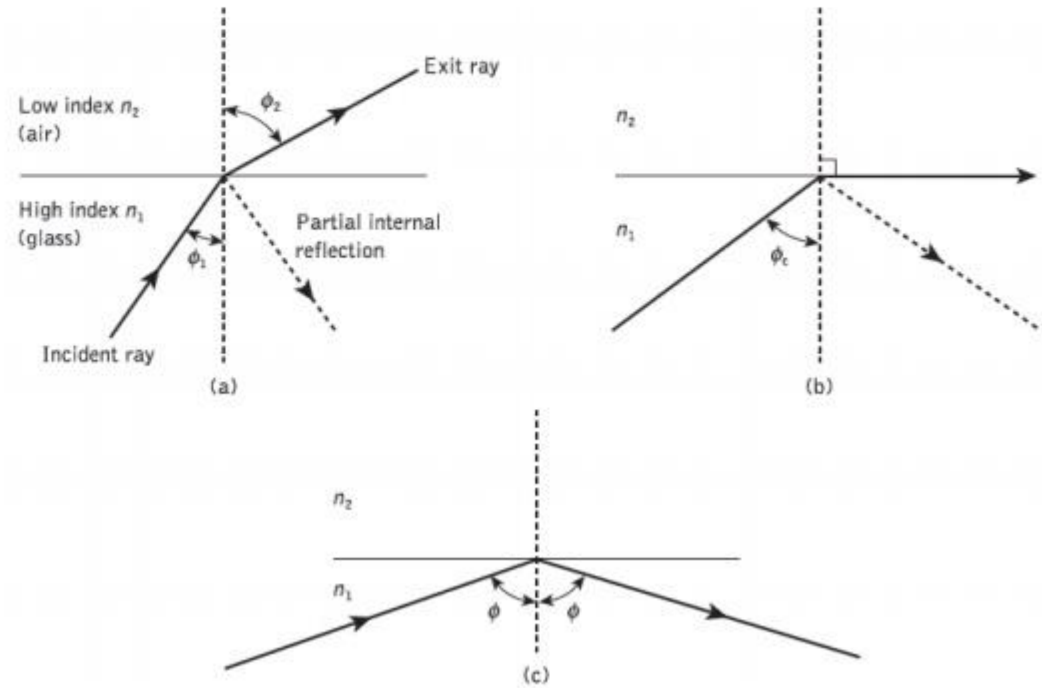
To consider the propagation of light within an optical fiber utilizing the ray theory model it is necessary to take account of the refractive index of the dielectric medium. The refractive index of a medium is defined as the ratio of the velocity of light in a vacuum to the velocity of light in the medium.

A ray of light travels more slowly in an optically dense medium than in one that is less dense, and the refractive index gives a measure of this effect. When a ray is incident on the interface between two dielectrics of differing refractive indices (e.g. glass–air), refraction occurs. It may be observed that the ray approaching the interface is propagating in a dielectric of refractive index *n* and is at an angle φ to the normal at the surface of the interface. If the dielectric on the other side of the interface has a refractive index *n* which is less than *n*1, then the refraction is such that the ray path in this lower index medium is at an angle to the normal, where is greater than . The angles of incidence and refraction are related to each other and to the refractive indices of the dielectrics by Snell's law of refraction, which states that:

$$n_1 \sin \phi_1 = n_2 \sin \phi_2$$

Or

$$\frac{\sin \phi_1}{\sin \phi_2} = \frac{n_2}{n_1}$$

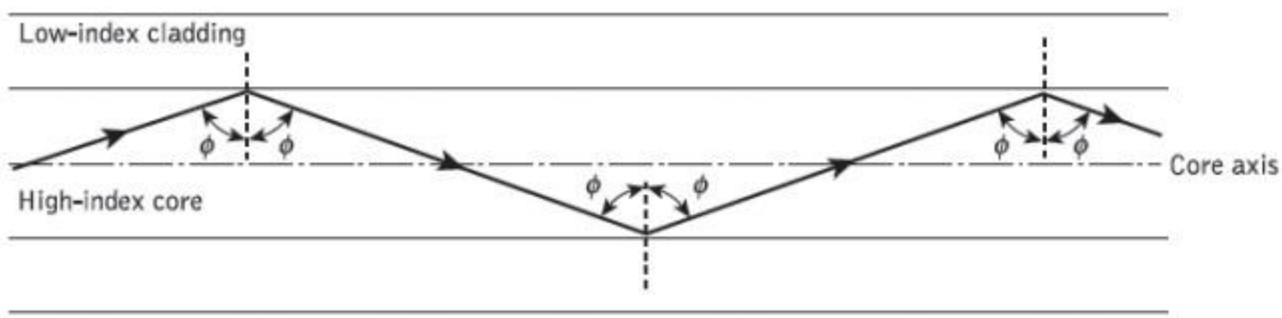


(a) Refraction, (b) Refraction showing the critical angle, (c) Total internal reflection

It may also be observed that a small amount of light is reflected back into the originating dielectric medium (partial internal reflection). As *n* is greater than *n*, the angle of refraction is always greater than the angle of incidence. Thus when the angle of refraction is 90° and the refracted ray emerges parallel to the interface between the dielectrics, the angle of incidence must be less than 90°. This is the limiting case of refraction and the angle of incidence is now known as the critical angle Φ_c, the value of the critical angle is given by

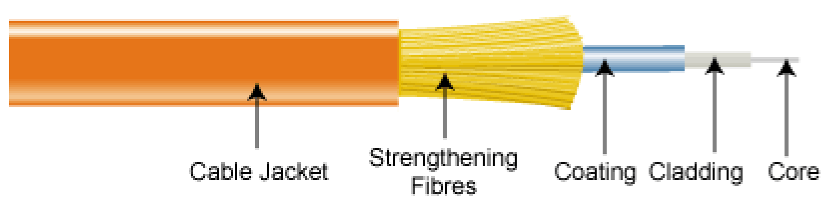
$$\sin \phi_c = \frac{n_2}{n_1}$$

At angles of incidence greater than the critical angle the light is reflected back into the originating dielectric medium (total internal reflection) with high efficiency (around 99.9%). Hence, it may be observed that total internal reflection occurs at the inter- face between two dielectrics of differing refractive indices when light is incident on the dielectric of lower index from the dielectric of higher index, and the angle of incidence of the ray exceeds the critical value. This is the mechanism by which light at a sufficiently shallow angle (less than 90° – may be considered to propagate down an optical fiber with low loss.



The above figure shows the transmission of a light ray in an optical fiber via a series of total internal reflections at the interface of the silica core and the slightly lower refractive index silica cladding. The ray has an angle of incidence Φ at the interface which is greater than the critical angle and is reflected at the same angle to the normal.

<u>FIBRE OPTIC CABLE CONSTRUCTION</u>



Core

This is the physical medium that transports optical data signals from an attached light source to a receiving device. The core is a single continuous strand of glass or plastic that's measured in microns (μ) by the size of its outer diameter. The larger the core, the more light the cable can carry.

All fiber optic cable is sized according to its core's outer diameter. The three multimode sizes most commonly available are 50, 62.5, and 100 microns. Single-mode cores are generally less than 9 microns.

Cladding

This is the thin layer that surrounds the fiber core and serves as a boundary that contains the light waves and causes the refraction, enabling data to travel throughout the length of the fiber segment.

Coating

This is a layer of plastic that surrounds the core and cladding to reinforce and protect the fiber core. Coatings are measured in microns and can range from 250 to 900 microns.

Strengthening fibers

These components help protect the core against crushing forces and excessive tension during installation. The materials can range from Kevlar to wire strands to gel-filled sleeves.

Cable jacket

This is the outer layer of any cable. Most fiber optic cables have an orange jacket, although some types can have black or yellow jackets.

VELOCITY OF PROPAGATION-

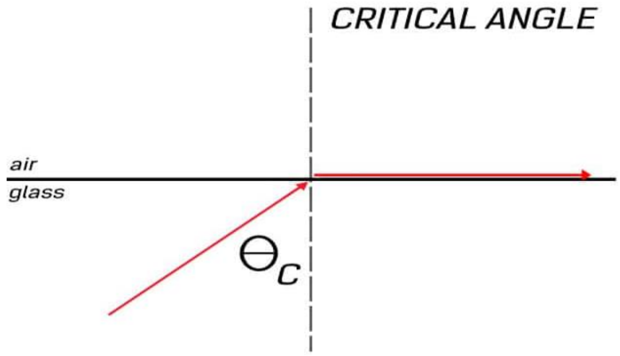
Velocity of propagation is a measure of how fast a signal travels over time, or the speed of the transmitted signal as compared to the speed of light.

In computer technology, the velocity of propagation of an electrical or electromagnetic signal is the speed of transmission through a physical medium such as a coaxial cable or optical fiber. There is also a direct relation between velocity of propagation and wavelength. Velocity of propagation is often stated either as a percentage of the speed of light or as time-to distance.

CRITICAL ANGLE-

When the angle of refraction is 90 degrees to the normal, the refracted ray is parallel to the interface between the two media. In this case, the incident angle is called the critical angle. You can calculate the critical angle using the following formula

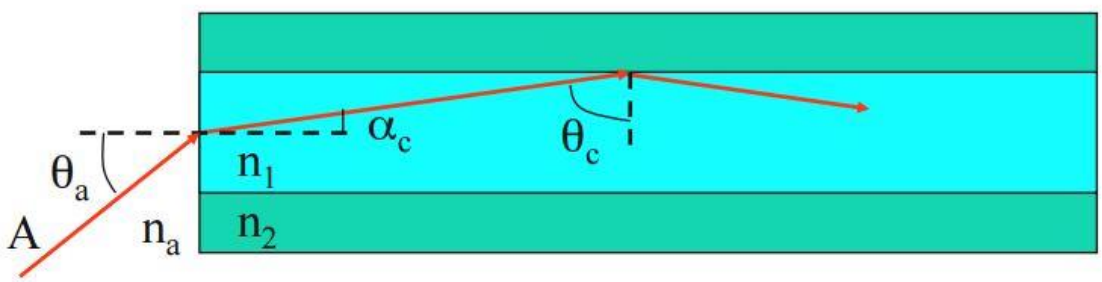
$$\theta_c = \sin^{-1} \frac{n_2}{n_1}$$



It is important to know about this property because reflection is also possible even if the surfaces are not reflective. If the *angle of incidence is greater than the critical angle* for a given setting, the resulting type of reflection is called **Total Internal Reflection**, and it is the basis of Optical Fiber Communication.

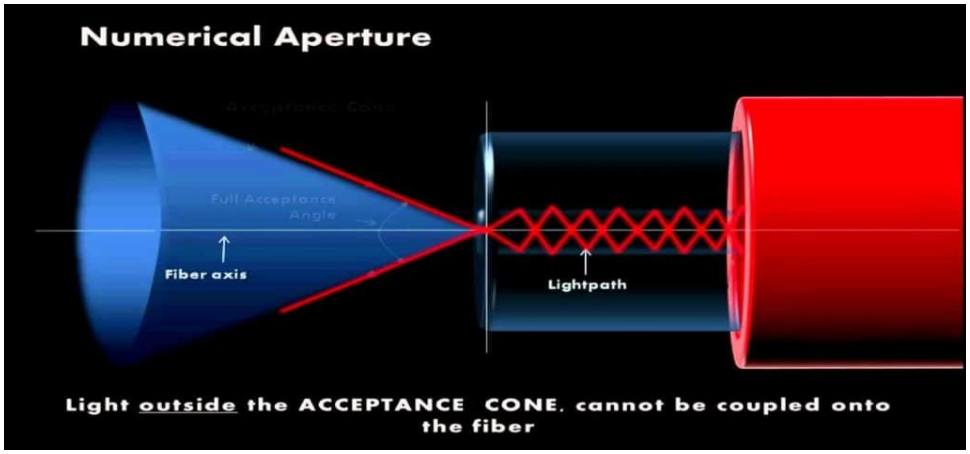
ACCEPTANCE ANGLE

In an optical fiber, a light ray undergoes its *first refraction* at the air-core interface. The angle at which this refraction occurs is crucial because this particular angle will dictate whether the subsequent *internal* reflections will follow the principle of Total Internal Reflection. This angle, at which the light ray first encounters the core of an optical fiber, is called the Acceptance angle.



NUMERICAL APERTURE

Numerical Aperture is a characteristic of any optical system. For example, photo-detector, optical fiber, lenses etc. are all optical systems. Numerical aperture is the ability of the optical system to collect all of the light incident on it, in one area.



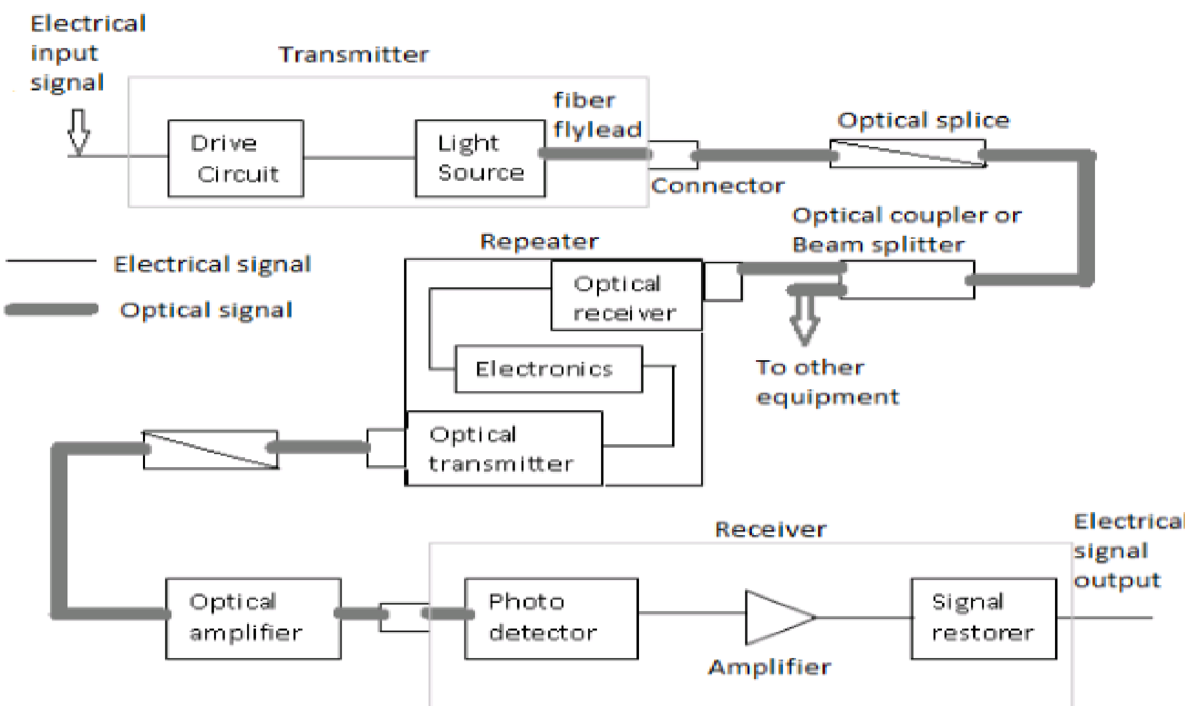
The blue cone is known as the cone of acceptance. As you can see, it is dependent on the Acceptance Angle of the optical fiber. Light waves within the acceptance cone can be collected in a small area which can then be sent into the optical fiber.

The numerical aperture of an optical fiber can be calculated with the following two formulas

$$NA = n_0 \sin \theta_a$$

$$NA = \sqrt{n_1^2 - n_2^2}$$

BLOCK DIAGRAM OF AN OPTICAL FIBER COMMUNICATION SYSTEM-



- The optical fiber consists of three main elements:

1. **Transmitter**: An electric signal is applied to the optical transmitter. The optical transmitter consists of driver circuit, light source and fiber fly lead.

- Driver circuit drives the light source.
- Light source converts electrical signal to optical signal.
- Fiber fly lead is used to connect optical signal to optical fiber.

2. **Transmission channel**: It consists of a cable that provides mechanical and environmental protection to the optical fibers contained inside. Each optical fiber acts as an individual channel.

- Optical splice is used to permanently join two individual optical fibers.
- Optical connector is for temporary non-fixed joints between two individual optical fibers.
- Optical coupler or splitter provides signal to other devices.
- Repeater converts the optical signal into electrical signal using optical receiver and passes it to electronic circuit where it is reshaped and amplified as it gets attenuated and distorted with increasing distance because of scattering, absorption and dispersion in waveguides, and this signal is then again converted into optical signal by the optical transmitter.

3. **Receiver**: Optical signal is applied to the optical receiver. It consists of photo detector, amplifier and signal restorer.

- Photo detector converts the optical signal to electrical signal.
- Signal restorers and amplifiers are used to improve signal to noise ratio of the signal as there are chances of noise to be introduced in the signal due to the use of photo detectors.

- For short distance communication only main elements are required.

  Source- LED

  Fiber- Multimode step index fiber

  Detector- PIN detector

- For long distance communication along with the main elements there is need for couplers, beam splitters, repeaters, optical amplifiers.

  Source- LASER diode

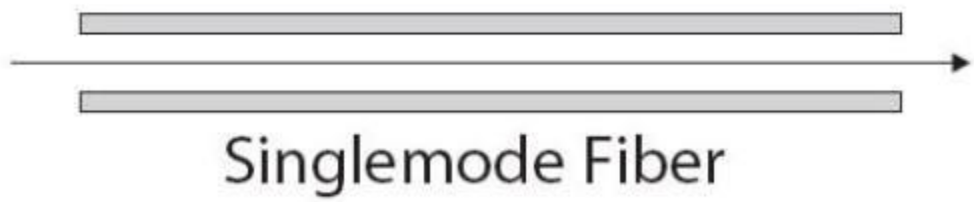
  Fiber- single mode fiber

  Detector- Avalanche photo diode (APD)

MODES OF PROPAGATION OF OPTICAL FIBER-

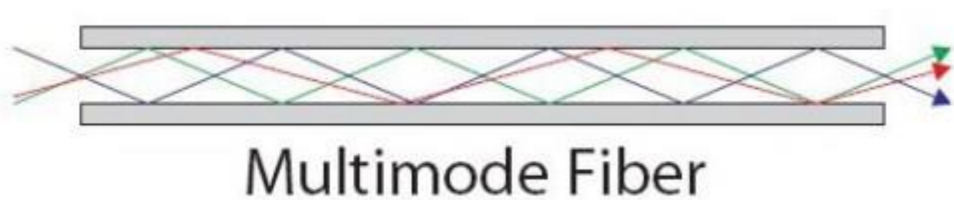
There are 2 types of propagation mode in fiber optics cable which are multi-mode and single-mode.

Single-mode optical fiber



- As the name suggests, this type of optical fiber transmits only one mode of light. To put it another way, it can carry only one wavelength of light across its length.
- This wavelength is usually 1310nm or 1550nm.
- You would think that this limits its capabilities of transferring more data. But single-mode type of optical fibers is much better than multimode optical fibers as they have more bandwidth and experience fewer losses. So the speed is unmatched.
- Interestingly, single-mode fibers came into existence *after* multimode fibers. They are more recent than the multimode cables.
- These cables can carry only one mode, physically, by having a tiny core. That is to say that the diameter of the core is essentially of the same order as the wavelength of the light passing through it.
- Only lasers are used as a light source. To point out, the light used in single-mode fibers are not in the visible spectrum.
- Since the light travels in a straight direction, there are fewer losses, and it can be used in applications requiring longer distance connections.
- A distinct disadvantage of single-mode fiber is that they are hard to couple.

Multimode optical fiber



- As the name implies, these types of optical fibers allow multiple modes of light to travel along their axis.
- To explain physically, they can do this by having a thicker core diameter.
- The wavelengths of light waves in multimode fibers are in the visible spectrum ranging from 850 to 1300 nm.
- The reflection of the waves inside the multimode fiber occurs at different angles for every mode. Consequently, based on these angles, the *number* of reflections can vary.

- We can have a mode where the light passes without striking the core at all.
- We can have a slightly higher mode, which will travel with appropriate internal reflections.
- Since the basis of optical fiber, communication is a total internal reflection, all modes with incident angles that do not cause total internal reflection get absorbed by the cladding. As a result, losses are created.
- We can have higher-order modes, waves that are highly transverse to the axis of the waveguide can reflect many times. In fact, due to increased reflections at unusual angles, higher-order modes can get completely lost inside the cable.
- Lower order modes are moderately transverse or even completely straight and hence fare better comparatively.

TYPES OF OPTICAL FIBER CONFIGURATION-

Single-mode Step Index

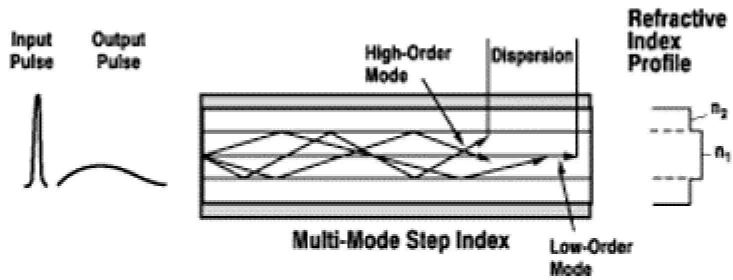
Single-Mode Step Index

The diameter of the core is fairly small relative to the cladding. Typically, the cladding is ten times thicker than the core. Comparing the output pulse and the input pulse note that there is little attenuation and time dispersion.

Single mode propagation exists only above a certain specific wavelength called the cutoff wavelength. Single-mode fiber optic cable is fabricated from glass. Because of the thickness of the core, plastic cannot be used to fabricate single-mode fiber optic cable.

Less time dispersion of course means higher bandwidth and this is in the 50 to 100 GHz/ km range. However, single mode fiber optic cable is also the most costly in the premises environment. For this reason, it has been used more with Wide Area Networks than with premises data communications. It is attractive more for link lengths go all the way up to 100 km. Nonetheless, single-mode fiber optic cable has been getting increased attention as Local Area Networks have been extended to greater distances over corporate campuses.

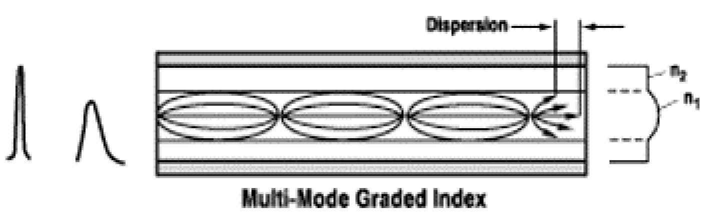
Multi-mode Step Index



The diameter of the core is fairly large relative to the cladding. Note that the output pulse is significantly attenuated relative to the input pulse. It also suffers significant time dispersion. The higher order modes, the bouncing rays, tend to leak into the cladding as they propagate down the fiber optic cable. They lose some of their energy into heat. This results in an attenuated output signal. Consequently, they do not all reach the right end of the fiber optic cable at the same time. When the output pulse is constructed from these separate ray components the result is time dispersion.

Fiber optic cable that exhibits multi-mode propagation with a step index profile is thereby characterized as having higher attenuation and more time dispersion than the other propagation candidates have. However, it is also the least costly and in the premises environment the most widely used. It is especially attractive for link lengths up to 5 km. usually, it has a core diameter that ranges from 100 microns to 970 microns. It can be fabricated either from glass, plastic or PCS.

Multi-mode Graded Index



There is no sharp discontinuity in the indices of refraction between core and cladding. The core here is much larger than in the single-mode step index. When comparing the output pulse and the input pulse, note that there is some attenuation and time dispersion, but not nearly as great as with multi-mode step index fiber optic cable.

Fiber optic cable that exhibits multi-mode propagation with a graded index profile is thereby characterized as having attenuation and time dispersion properties somewhere between the other two candidates. Likewise its cost is somewhere between the other two candidates. This type of fiber optic cable is extremely popular in premise data communications applications.

LOSSES IN OPTICAL FIBERS-

ATTENUATION -

- Attenuation is the loss of optical energy as it travels through the fiber, this loss of optical energy as it travels through the fiber, this loss is measured in dB.
- Attenuation is a transmission loss that can be measured as a difference between the output signal power and the input signal power.

Attenuation loss= α = 10 log_{10} (P_{input}/P_{out})dB

- Attenuation depends on-
  1) Attenuation depends on wavelength used.
  2) Attenuation depends on light intensity i.e. input light power.
  3) It depends on diameter of optical fiber.
  4) It depends on distance. Distance between optical source and repeater/detector.

Different types of losses in optical fiber.

1. ABSORPTION-

It is the way by which the energy of a photon is taken up by matter. Thus the light energy is transformed to other forms of energy i.e, to heat. The absorption of light during wave propagation is called attenuation.

Absorption of light in optical fibers may be intrinsic or extrinsic.

Intrinsic Absorption-

Intrinsic absorption is caused by basic fiber material properties. Intrinsic absorption is due to material and electron absorption.

Material absorption is a loss mechanism which results in the dissipation of some of the transmitted optical power into heat in the optical fiber.

Intrinsic absorption in the ultraviolet region is caused by electronic absorption. Basically, absorption occurs when a light particle interacts with an electron and excites it to higher energy level.

Extrinsic Absorption-

Extrinsic absorption is caused by impurities introduced into the fiber material. Metal impurities such as, iron, nickel and chromium are introduced into the fiber during fabrication.

2. SCATTERING LOSS-

Scattering loss occurs when light strikes a substance which emits light of its own at the same wavelength as the incident light. The propagation of light through the core of an optical fiber is based on total internal reflection of the light wave. Rough and irregular surfaces can cause light rays to be reflected in random directions. This is called scattering.

3. **BENDING LOSSES-**

The loss which exists when an optical fiber undergoes bending is called bending losses. There are two types of bending-

(b) Macroscopic Bending-

Bending in which complete fiber under goes bends which causes certain modes not to be reflected and therefore causes loss to the cladding.

(c) Microscopic Bending-

Either the core or cladding undergoes slight bends at its surface. It causes light to be reflected at angles when there is no further reflection

4. DISPERSION-

Dispersion is the spreading out of a light pulse in time as it propagates down the fiber. Dispersion is of 2 types.

- Intermodal Dispersion
- Intra modal Dispersion

Intermodal Dispersion-

Multimode fibers can guide many different light modes since they have much larger core size. Each mode enters the fiber at a different angle and thus travels at different paths in the fiber. Since each mode ray travels at different distance as it propagates, the ray arrive at different times at the fiber output.

So, the light pulse spreads out in time which can cause signal overlapping. Intermodal dispersion is not a problem in single mode fibers since there is only one mode.

Intra modal Dispersion-

There are two types of intra modal Dispersion.

- Material Dispersion
- Waveguide Dispersion

Material Dispersion-

Material dispersion is a type of chromatic dispersion. Chromatic dispersion is the pulse spreading that arises because the velocity of light through a fiber depends on its wavelength.

Waveguide Dispersion-

It is caused by the fact that some light travels in the fiber cladding compared to most light travels in the fiber core. Since fiber cladding has lower index than fiber core, light ray that travels in the cladding travels faster than that in the core.

OPTICAL SOURCES AND ITS TYPE

Like other communication system, fiber optic communication has also a transmitter at one end of the system, which injects information on to the fiber cables. The transmitter processes and translates coded electronic pulse information coming from copper wire into equivalently coded light pulses. The basic concept behind the optical transmitter is that it converts electrical input signals into modulated light for transmission over an optical fiber. The input signal determines the characteristics of the resulting modulated light, which may be turned on and off or may be linearly varied in intensity between two predetermined levels.

Here are two commonly· used optical sources for generating the light pulses. These are light emitting diode (LED) and Laser Diode (LD). Laser diode with its version as injection-laser diode (ILD) is commonly employed. Both the sources funnel the light pulses into the fiber-optic medium where they transmit themselves down the fiber cable and are placed in very close proximity to the light emitting region to couple as much light as possible into the fiber. To accomplish the same, they are mounted in a package that enables an optical fiber to couple as .much light as possible into the fiber.

Concept of LED-

An LED or a Light Emitting Diode is semiconductor device that emits light due to Electroluminescence effect. An LED is basically a PN Junction Diode, which emits light when forward biased. Light Emitting Diodes are almost everywhere. You can find LEDs in Cars, Bikes, Street Lights, Home Lighting, Office Lighting, Mobile Phones, Televisions and many more.

The reason for such wide range of implementation of LEDs is its advantages over traditional incandescent bulbs and the recent compact fluorescent lamps (CFL). Few advantages of LEDs over incandescent and CFL light sources are mentioned below:

- Low Power Consumption
- Small Size
- Fast Switching
- Physically Robust
- Long Lasting

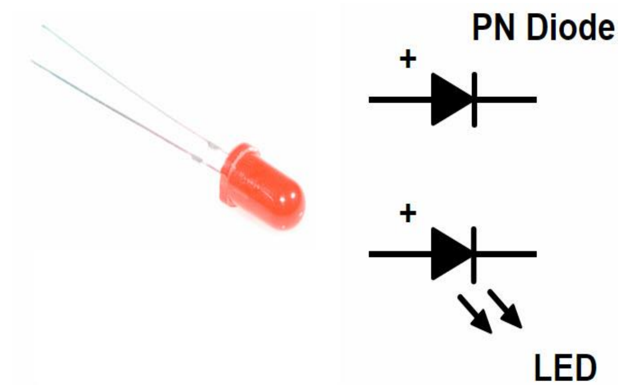
Because of these advantages, LEDs have become quite popular among a large set of people. Electronics Engineers, Electronic Hobbyists and Electronics Enthusiasts often work with LEDs for various projects.

It consists of a PN Junction Diode and when voltage is applied to the LED, electrons and holes recombine in the PN Junction and release energy in the form of light (Photons).

The light emitted by an LED is usually monochromatic i.e. of single color and the color is dependent on the energy band gap of the semiconductor.

Light Emitting Diodes can be manufactured to emit all the wavelengths of visible spectrum i.e. from Red (620nm to 750nm) to blue – violet (380nm to 490nm).

The electrical symbol of an LED is similar to that of a PN Junction Diode. The following image shows a Red LED along with symbols of PN Junction Diode and LED.



Types of LED-

There are two basic types of LED structures:

1. Surface Emitting LED
2. Edge Emitting LED

Surface Emitting LED-

• It is modified form of DH LED (Double Heterojunction LED). In this LED type, optical fiber is butt-coupled with itself.

• The surface emitting LED structure consists of thin central active layer of p type GaAs.

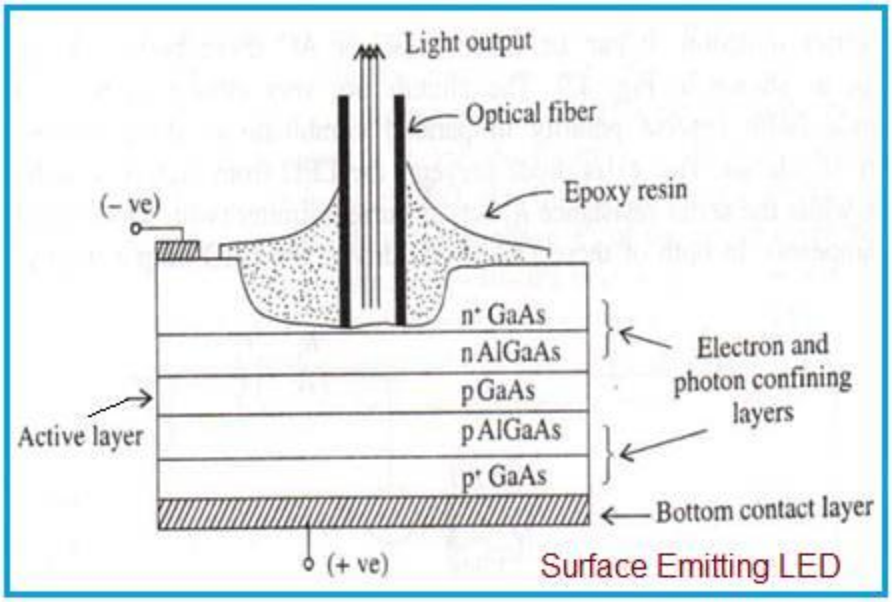
• This central layer is bounded by n-type AlGaAs/n^+-type GaAs at the top side.

• This central layer is bounded by p-type AlGaAs/p^+-type GaAs at the bottom side.

• The extreme top n^+ type GaAs and bottom p^+-type GaAs layers are used to provide low resistive ohmic contacts only.

• The external optical fiber is butt connected by etching the top layers and by shielding with epoxy resin.

• When refractive indices of both p-type and n-type materials are same, light is free to come out from all the sides of the semiconductor device due to no confinement. However only active region near the surface will emit the significant amount of light while absorbing from the other parts. Hence it is known as surface emitting LED.



• Output radiation is originated from central thin layer i.e. p-type GaAs layer.

Advantages of Surface Emitting LED:

1) Optical coupling coefficient of LED with external fiber system is relatively higher. Hence this LED offers high optical coupling efficiency.

2) Optical loss (due to internal absorption) is very low. This is because of carrier recombination near its top heterojunction.

3) InP/InGaAsP based LED is used for long wavelength applications.

4) It offers higher efficiency with low to high radiance.

5) The top n-GaAs contact layer ensures low thermal resistance and contact resistance. This allows high current densities and high radiation intensity.

6) The internal absorption in the device is very low due to larger bandgap confining layers. Moreover reflection coefficient at the back crystal face is high which gives good forward radiance.

Disadvantages of Surface Emitting LED:

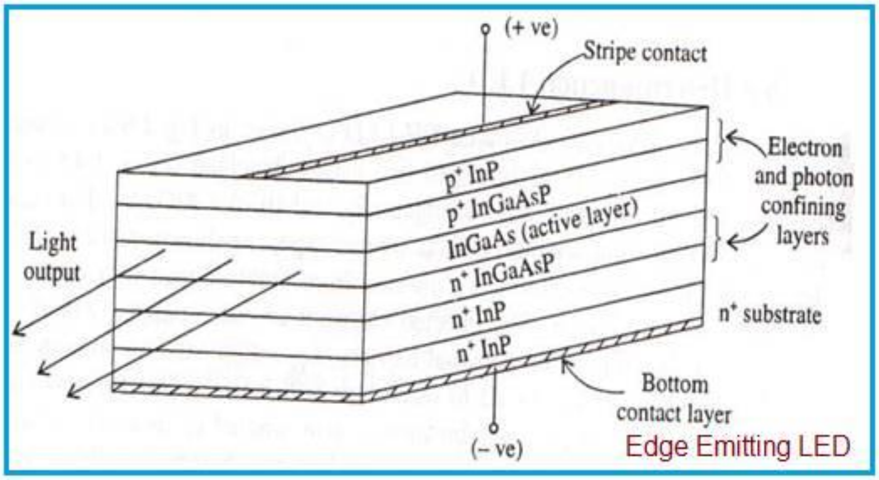
1) The surface emitting LED can transmit data rate less than 20 Mbps than edge emitting LED.

2) It contains short optical link with large NA (Numerical Aperture).

<u>Edge Emitting LED-</u>

• It is widely used in optical fiber communication system. Here collimated light from LED is required to be fed into the fiber with high coupling efficiency.

• It is used for long wavelength optical communication approx. between 1.33 to 1.55 μm.



• Modern epitaxial growth techniques such as MBE, MOCVD etc. are used in order to design such complex LED structures.

• Central active layer is made using InGaAs having narrow bandgap. It is bounded by wide bandgap layers such as p+ InGaAsP and n+ InP cladding layers.

• These two cladding layers help in confining injected electrons and holes into the middle layer. It also helps emitted photons to travel along LED axis as per optical properties.

• Due to above, light is emitted from the edge of the LED. Hence it is known by the name edge emitting LED.

Advantages of Edge Emitting LED:

1) It has superior beam collimation property which offers greater coupling efficiency with fiber optic cable compare to surface emitting LED.

2) It offers higher efficiency with low to high radiance. It fulfills high brightness LED requirements of the lighting industry.

3) It radiates less power to the air compare to surface emitting LED due to reabsorption and

interfacial recombination.

4) It offers better modulation bandwidth and more directional emission pattern.

5) It offers 5-6 times more coupled power into NA (Numerical Aperture) of step/graded index fibers. This is due to small beam divergence.

6) It offers high data rates (> 20 Mbps) than surface emitting LED.

Disadvantages of Edge Emitting LED:

1) Its structure is complex.

2) It is difficult to design heat sink.

3) It is expensive compare to other LED types.

4) There are many issues to be handled during mechanical mounting and installation.

DEFINE LASER AND ITS WORKING PRINCIPLE-

A LASER diode is an optoelectronic device, which converts electrical energy to produce high intensity coherent light. In a Laser diode, the P-N junction of the semiconductor diode acts as the active medium.

Construction-

The Laser diode is made of two doped gallium arsenide layers. One doped gallium arsenide layer will produce an n-type semiconductor whereas another doped gallium arsenide layer will produce a p-type semiconductor. In Laser diodes, selenium, aluminium and silicon are used as doping agents.

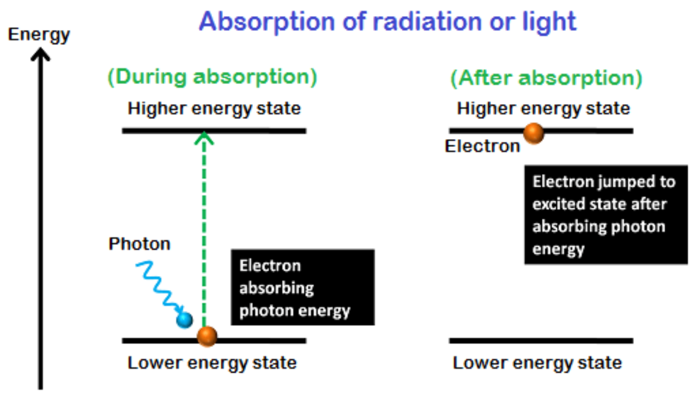
Working-

In lasers, photons are interacted in three ways with the atoms:

- Absorption of radiation
- Spontaneous emission
- Stimulated emission

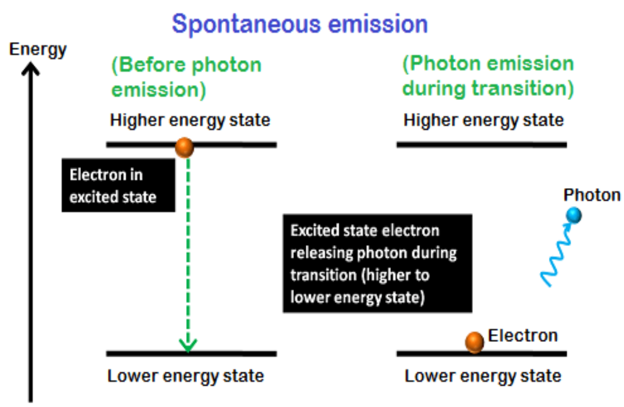
Absorption of radiation

Absorption of energy is the process of absorbing energy from the external energy sources.



Spontaneous emission

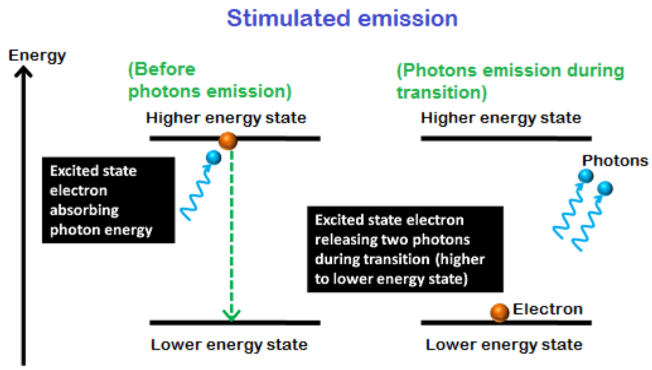
Spontaneous emission is the process of emitting light or photons naturally while electrons falling to the lower energy state.



Stimulated emission

Stimulated emission is the process by which excited electrons are stimulated to fall into the lower energy state by releasing energy in the form of light. The stimulated emission is an artificial process.

In spontaneous emission, the electrons in the excited state will remain there until its lifetime is over. After completing their lifetime, they return to the ground state by releasing energy in the form of light.



- When DC voltage is applied across the laser diode, the free electrons move across the junction region from N-type material to the P-type material.
- In this process some electron will directly interact with the valence electrons and excites them to the higher energy level whereas some other electrons will recombine with the holes in the P-type semiconductor and releases energy in the form of light. This process of emission is called spontaneous emission.
- The photons generated due to spontaneous emission will travel through the junction region and stimulate the excited electrons.
- As a result, more photons are released. This process of light or photons emission is called stimulated emission will moves parallel to the junction.
- The two ends of the laser diode structure are optically reflective. One end is fully reflective end will reflect the light completely whereas the partially reflective end will reflect most part of the light but allows a small amount of light.
- The light generated due to the stimulated emission is escaped through the partially reflective end of the laser to produce a narrow beam laser light.

Advantages-

1) Simple in construction.
2) Light weight.
3) Small in size.
4) High efficiency.
5) Low power consumption.

Disadvantages-

1) Not suitable for the applications where high powers are required.
2) Semiconductor lasers are highly dependent on temperature.

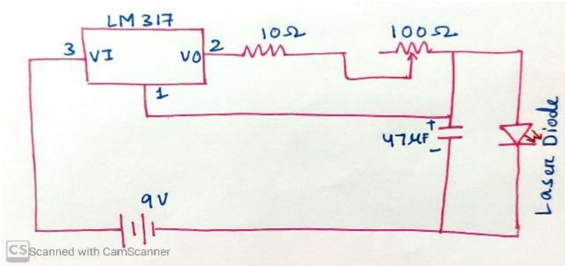
Application-

1) Laser diodes are used in laser pointer.
2) It is used in fiber optic cable.
3) It is used in laser scanning.

<u>LASER FEEDBACK CONTROL CIRCUIT-</u>

- LASER diodes require complex drive circuit that involve feedback loops by measuring output optical power, temperature voltage and input current.

- But for controlling a laser diode used in applications where high accuracy is not required, a simple LASER diode driver circuit can be constructed using LM 317 voltage regulator IC.



- The LM 317 is configured to function as a constant current source. The output current depends on the value of resistance between V_{out} and V_{adj} of LM 317.
- So adjusting the 100 `Ω potentiometer will change the output current that flows into the LASER diode. The 10`Ω resistor is used to prevent large currents from flowing when the value of potentiometer is at zero. The 47μF capacitor is used to absorb any battery voltage spikes.

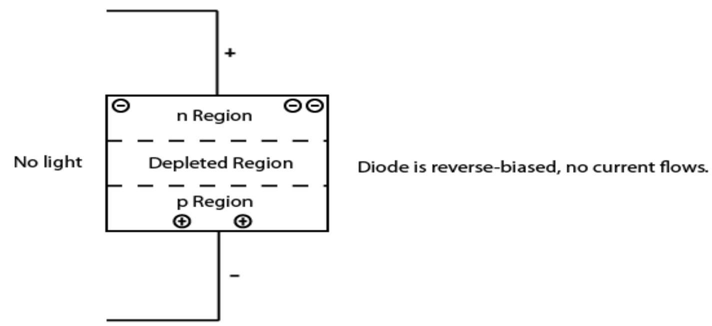
OPTICAL DETECTOR-

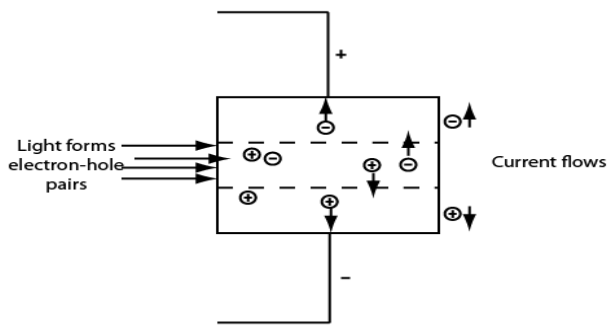
An optical detector is a device that converts light signals into electrical signals, which can then be amplified and processed. The photodetector is as essential an element of any fiber optic system as the optical fiber or the light source. Photodetectors can dictate the performance of a fiber optic communication link.

WORKING OF PIN DIODE-

PIN photodetector has an intrinsic semiconductor region sandwiched between a P-doped and an n-doped region. Since the intrinsic region has no free charges, its resistance is high, so that most of the reverse biased voltage is applied to this i region. The i region is usually wide so that incoming photons have a greater probability of absorption in the i region rather than in the p or n regions. Since the electric field is high in the i region, any electron hole pairs generated in this region are immediately swept away by the field.

The detector is electrically reverse-biased. (In contrary, LEDs and Lasers are forward-biased to emit light).





In the first illustration when there is no light, the reverse bias draws current-carrying electrons and holes out of the p-n junction region, creating a depleted region, which stops current from passing through the diode.

In the second illustration when there are lights on the detector, photons with the proper energy (wavelength) can create electron-hole pairs in this region by raising an electron from the valence band to the conduction band, leaving a hole behind. The bias voltage causes these current carriers to drift quickly away from the junction region, so a current flows proportional to the light hitting the detector.

AVALANCHE PHOTODIODE-

- An avalanche photodiode is a semiconductor-based **photodetector** (**photodiode**) which is operated with a relatively high reverse voltage (typically tens or even hundreds of volts), sometimes just below breakdown.
- In this region, carriers (electrons and holes) excited by absorbed **photons** are strongly accelerated in the strong internal electric field, so that they can generate secondary carriers.
- The avalanche process, which may take place over a distance of only a few micrometers, for example, effectively amplifies the **photocurrent** by a significant factor, although not as much as in a **photomultiplier**.
- Therefore, avalanche photodiodes can be used for very sensitive detectors, which need less electronic signal amplification and are thus less susceptible to electronic noise.

- However, the avalanche process itself is subject to **quantum noise** and amplification noise, which can offset the mentioned advantage. The excess noise is quantified with the excess noise factor *F*, which is the factor by which the electronic noise power is increased compared with that of an ideal photodetector.
- Note that the amplification factor and thus the effective **responsivity** of an APD depends strongly on the reverse voltage, and may also substantially vary from device to device.

- Generally, the noise performance of photodetectors with APDs can be better than that of devices with ordinary **p–i–n photodiodes** when electronic noise is a limiting factor: the internal amplification in an APD reduces the influences of electronic noise. However, the above mentioned excess noise factor increases with increasing amplification factor, as obtained for increasing reverse voltage.
- Therefore, the reverse voltage is often chosen such that the multiplication noise approximately equals the noise of the electronic amplifier, because this minimizes the overall noise. The amount of excess noise depends on many factors: the magnitude of the reverse voltage, material properties (in particular, the ionization coefficient ratio), and the device design.
- **Advantages-**
  1. High level of sensitivity as a result of avalanche gain.
- **Disadvantages-**
  1. Much higher operating voltage may be required.
  2. Avalanche photodiode produces a much higher level of noise than a PIN photodiode.
  3. Avalanche process means that the output is not linear.
- The avalanche photodiodes are not as widely used as their PIN counterparts. They are used primarily where the level of gain is of paramount importance, because the high voltages required, combined with a lower reliability means that they are often less convenient to use.

OPTICAL CONNECTORS-

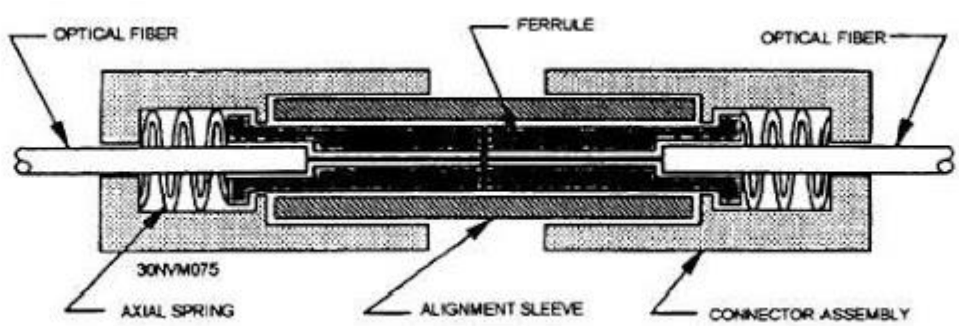
An optical fiber connector is a flexible device that connects fiber cables requiring a quick connection and disconnection.

Types of connector-

1. Butt- joint connector
2. Expanded beam connector

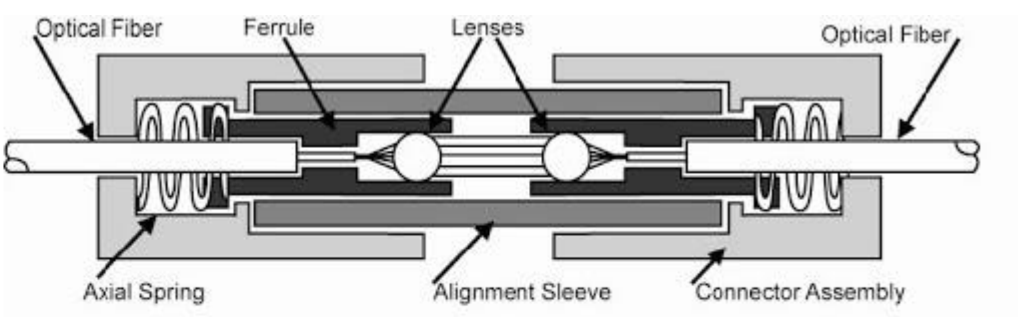
Butt-joint connector-

Butt joint connectors employ a metal, ceramic or molded plastic ferrule for each fiber and a precision sleeve into which the ferrule fit. The fiber is epoxied into a precision hole which has been drilled into the ferrule.



Expanded beam connector-

Expanded beam connector, employs lenses on the ends of the fibers. These lenses either collimate the light emerging from the transmitting fiber, or focus the expanded beam onto the core of the receiving fiber. The fiber to lens distance is equal to the focal length of the lens.



SPLICING-

A fiber optic splice is defined by the fact that it gives a permanent or relatively permanent connection between two fiber optic cables.

Types of Splicing

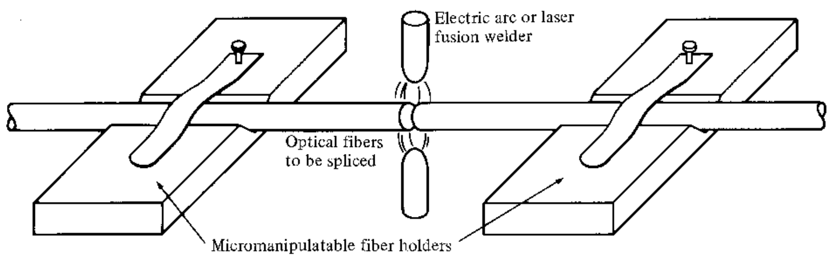
There are two main types of splicing

i) Fusion splicing.

ii) Mechanical splicing / V groove

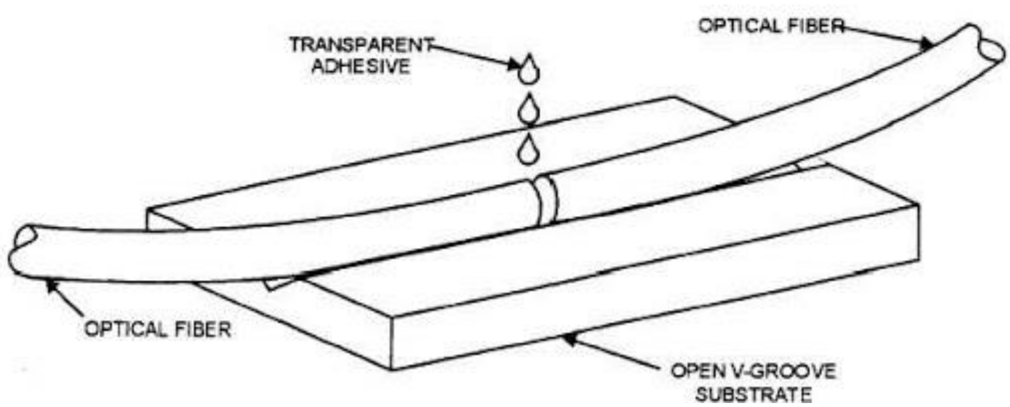
Fusion Splicing

Fusion splicing involves butting two cleaned fiber end faces and heating them until they melt together or fuse. Fusion splicing is normally done with a fusion splicer that controls the alignment of the two fibers to keep losses as low as 0.05 dB. Fiber ends are first pre aligned and butted together under a microscope with micromanipulators. The butted joint is heated with electric arc or laser pulse to melt the fiber ends so can be bonded together.



Mechanical Splicing / V Groove

Mechanical splices join two fibers together by clamping them with a structure or by epoxying the fibers together. Mechanical splices may have a slightly higher loss and back reflection. These can be reduced by inserting index matching gel. V groove mechanical splicing provides a temporary joint i.e fibers can be disassembled if required. The fiber ends are butted together in a V – shaped groove.



OPTICAL COUPLER-

A fiber optic coupler is an optical device capable of connecting one or more fiber ends in order to allow the transmission of light waves in multiple paths. The device is capable of combining two or more inputs into a single output and also dividing a single input into two or more outputs. Compared to a splice or connector, the signal can be more attenuated by fiber optic couplers, as the input signal can be divided amongst the output ports.

Passive and Active Couplers

Fiber optic couplers can either be passive or active devices.

Passive fiber optic couplers are said to be passive as no power is required for operation. They are simple fiber optic components that are used to redirect light waves. Passive couplers either use micro-lenses, graded-refractive-index (GRIN) rods and beam splitters, optical mixers, or splice and fuse the core of the optical fibers together.

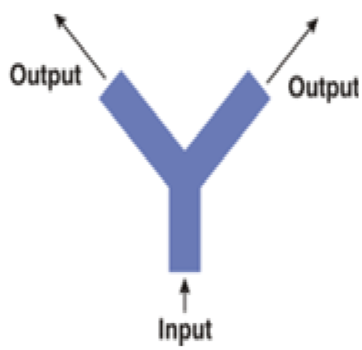
Active fiber optic couplers require an external power source. They receive input signal(s), and then use a combination of fiber optic detectors, optical-to-electrical converters, and light sources to transmit fiber optic signals.

TYPES

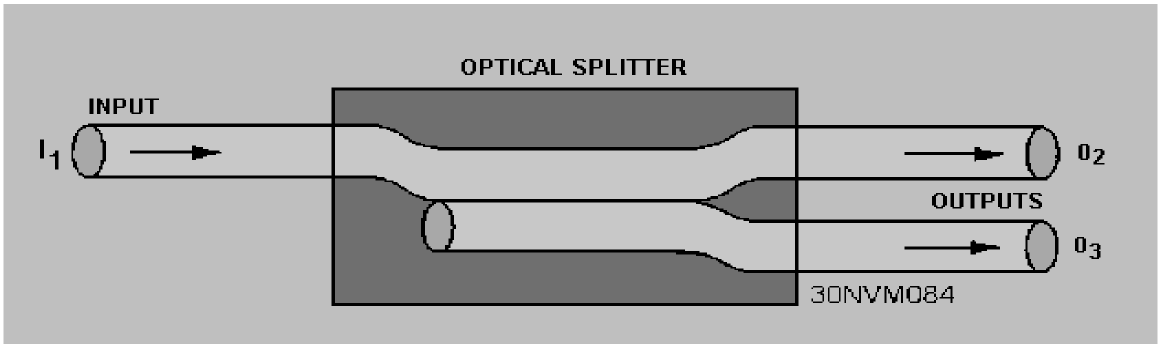
Types of fiber optic couplers include splitters, combiners, X-couplers, trees, and stars, which all include single window, dual window, or wideband transmissions.

Fiber optic splitters take an optical signal and supply two outputs. They can further be described as either Y-couplers or T-couplers.

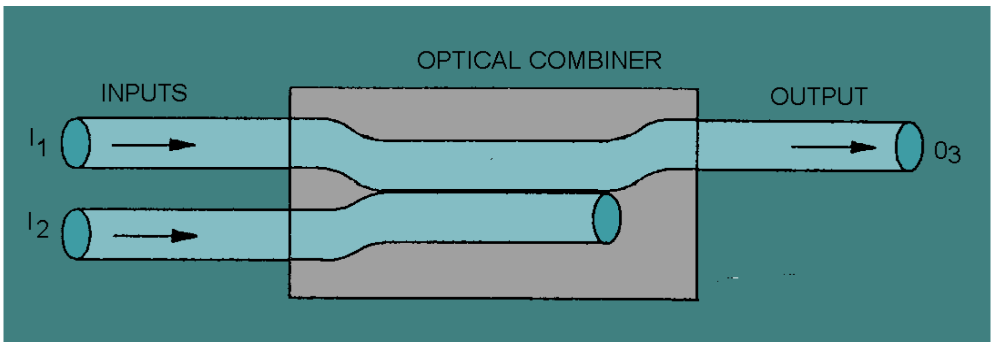
- **Y-couplers** have equal power distribution, meaning that the two output signal each receive half of the transmitted power.



- **T-couplers** have an uneven power distribution. The signal outputs still carrier the same signal, however the power of one output is greater than the second output.

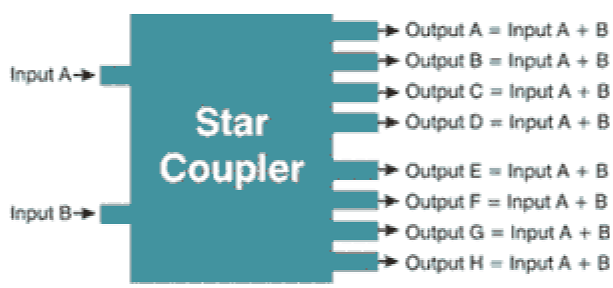


Fiber optic combiners receive two signals and provide a single output. The output signal is typically comprised of multiple wavelengths, due to the amount of interference that occurs when attempting to combine two signals that share the same wavelength.

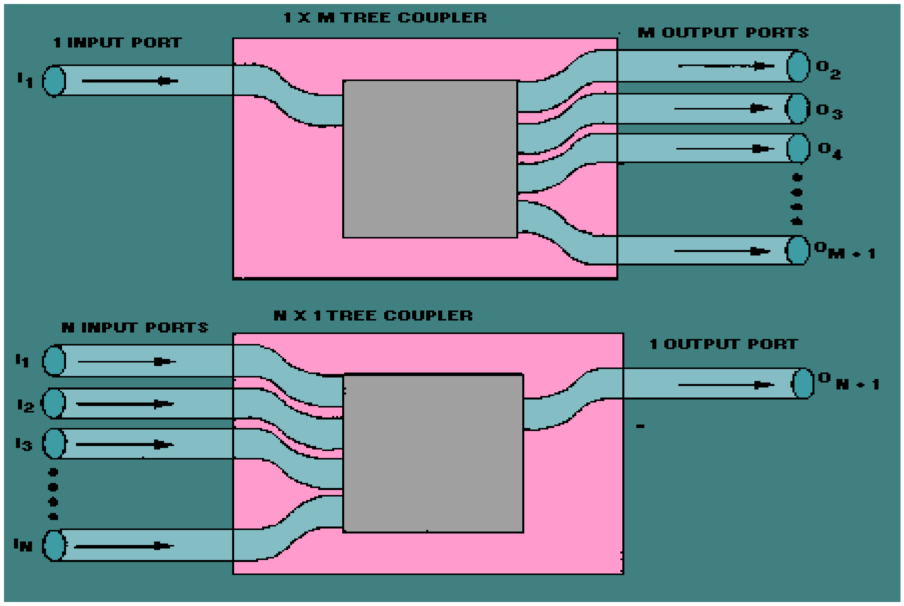


X-couplers carry out the function of a splitter and combiner in one package. They are a 2x2 coupler that combines the power of two signals and then divides the signal between two outputs.

Star couplers have M inputs and N outputs (MxN). They are used to distribute the power from all of the inputs to all outputs.



Tee couplers either have 1 input and M outputs (1xM) or N inputs and 1 output (Nx1).



OPTICAL REPEATER

An optical communications repeater is used in a fiber-optic communications system to regenerate an optical signal. Such repeaters are used to extend the reach of optical communications links by overcoming loss due to attenuation of the optical fiber. Some repeaters also correct for distortion of the optical signal by converting it to an electrical signal, processing that electrical signal and then retransmitting an optical signal. Such repeaters are known as optical-electrical-optical (OEO) due to the conversion of the signal. These repeaters are also called regenerators.

Optical regenerations are classified into 3 categories by the 3 R's scheme.

1. R: re-amplification of the data pulse alone is carried out.
2. 2R: in addition to re-amplification, pulse reshaping is carried out.
3. 3R: in addition to re-amplification and reshaping, retiming of data pulse is done.

There are two types of Fiber Repeater.

SFP to SFP Fiber Mode Repeaters: To simply interconnect two fiber connections back to back, an SFP to SFP Fiber Mode Repeater is generally used.

Ethernet Fiber Mode Repeater: In an environment where Ethernet is exclusively used, choosing a Media Converter or Repeater designed specifically for Ethernet fiber applications is recommended. They have Ethernet transceivers that regenerate the signal received before passing it along to the other fiber transceiver port. 3R (Re-amplify, Reshape, and Retime) signal regeneration ensures a strong signal across each fiber link to ensure link and data transmission integrity.

APPLICATION OF OPTICAL FIBER-

The application and uses of optical fiber can be seen in:

- Medical Industry
- Communication
- Defense
- Industries
- Broadcasting
- Lighting and Decorations
- Mechanical Inspections

The application of optical fibers in various fields are given below:

1**. Optical Fibers uses in Medical industry**

Because of the extremely thin and flexible nature, it used in various instruments to view internal body parts by inserting into hollow spaces in the body. It is used as lasers during surgeries, endoscopy, microscopy and biomedical research.

2**. Optical Fibers used in Communication**

In the communication system, telecommunication has major uses of optical fiber cables for transmitting and receiving purposes. It is used in various networking fields and even increases the speed and accuracy of the transmitted data. Compared to copper wires, fiber optics cables are lighter, more flexible and carry more data.

3. **Optical Fibers used in Defense Purpose**

Fiber optics are used for data transmission in high level data security fields of military and aerospace applications. These are used in wirings in aircrafts, hydrophones for SONARs and Seismic applications.

4. **Optical Fibers are used in Industries**

These fibers are used for imaging in hard to reach places such as they are used for safety measures and lighting purposes in automobiles both in the interior and exterior. They transmit information in lightning speed and are used in airbags and traction control. They are also used for research and testing purposes in industries.

5**. Optical Fibers used for Broadcasting**

These cables are used to transmit high definition television signals which has a greater bandwidth and speed. Optical Fiber is cheaper compared to same quantity of copper wires. Broadcasting companies use optical fibers for wiring HDTV, CATV, video-on demand and many applications.

6. **Uses of Optical Fiber for Lightening and Decorations**

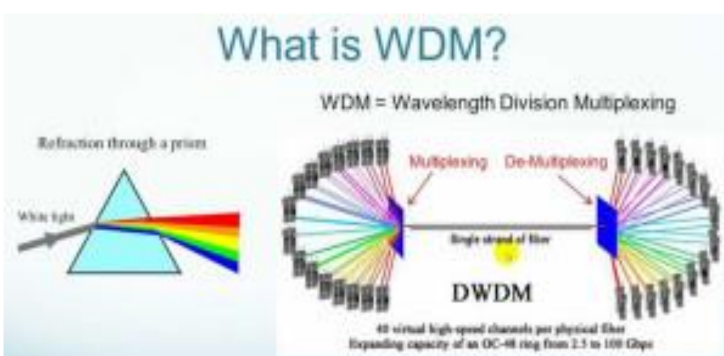
By now, we got a fair idea of what is optical fiber and it also gives an attractive, economical and easy way to illuminate the area and that is why, it is widely used in decorations and Christmas trees.

7. **Optical Fibers used in Mechanical Inspections**

On-site inspection engineers use optical fibers to detect damages and faults which are at hard to reach places. Even plumbers use optical fibers for inspection of pipes.

WAVELENGTH DIVISION MULTIPLEXING (WDM)-

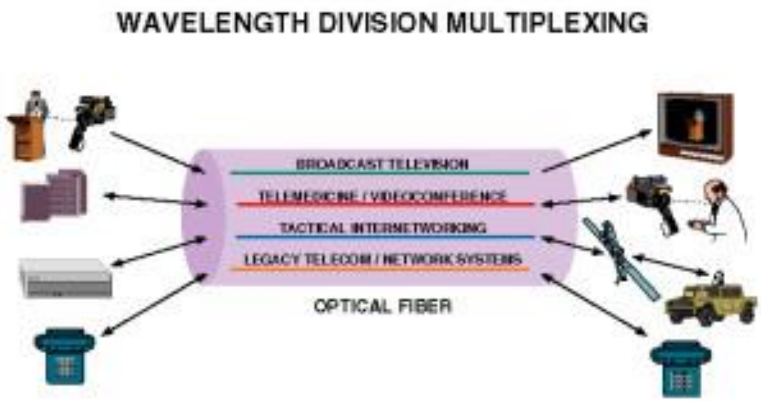
Wavelength division multiplexing (WDM) is a technique modulating various data streams, i.e. optical carrier signals of varying wavelengths in terms of colors of laser light onto a single optical fiber. Wavelength division multiplexing WDM is similar to frequency-division multiplexing (FDM) but referencing the wavelength of light to the frequency of light. WDM is done in the IR portion of the electromagnetic spectrum instead of taking place at radio frequencies (RF). Each IR channel carries several RF signals combined with frequency-division multiplexing (FDM) or time-division multiplexing (TDM). Each multiplexed infrared channel is separated or de-multiplexed into the original signals at final point. Data in different formats and at different speeds can be transmitted simultaneously on a single fiber by using FDM or TDM in each IR channel in combination with WDM. It allows network capacity to be gradually and cost effectively increased.



Wavelength Division Multiplexing (WDM)

What is Wavelength Division Multiplexing?

- WDM enables bi-directional communication and multiplies signal capacity. Each laser beam is modulated by separate set of signals. Since wavelength and frequency have an inverse relationship (shorter wavelength means higher frequency), the WDM and FDM both contains the same technology in them.
- At the receiving end, Wavelength-sensitive filters, IR analog of visible-light color filters are used.
- The first WDM systems were two-channel systems that used 1310nm and 1550nm wavelengths.



- Wavelength division multiplexing systems can combine signals with multiplexing and split them apart with a de-multiplexer. WDM systems are popular with telecommunications companies because they allow them to expand the capacity of the network without laying more fiber by using WDM and optical amplifiers.
- These two devices work as drop multiplexer (ADM), i.e. simultaneously adding light beams while dropping other light beams and rerouting them to other destinations and devices and this type of filtering of light beams were made possible with e talons, devices called Fabry-Perot interferometers using thin-film-coated optical glass.
- Present modern systems can handle up to128 signals and can expand a basic 9.6 Gbps fibre system to a capacity of over 1000 Gbps. It is mostly used for optical fiber communications to transmit data in several channels with slight variation in wavelengths. WDM can increase the total bit rate of point-to-point systems.

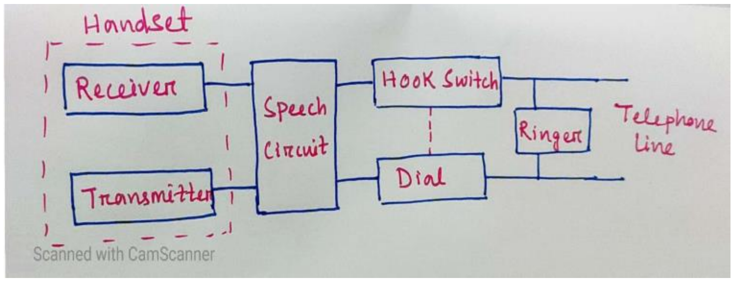
Uses of Wavelength Division Multiplexing:

- WDM multiply the effective bandwidth of a fiber optic communications system
- A fiber optic repeater device called the erbium amplifier can make WDM a cost-effective and it is the long-term solution.

- This reduces the cost and increases the capacity of the cable to carry data.
- Wavelength Division Multiplexing (WDM) uses multiple wavelengths (colors of light) to transport signals over a single fiber.
- It uses light of different colours to create a number of signal paths.
- It uses Optical prisms to separate the different colours at the receiving end and optical prisms does not require power source.
- These systems used temperature stabilized lasers to provide the needed channels count.

UNIT-4: TELECOMMUNICATION SYSTEM

BLOCK DIAGRAM OF TELEPHONE SET-



The telephone set consists of the following parts:

1. Microphone
2. Receiver
3. Switch connections to the telephone system
4. Ringing circuitry
5. Dial network

- The instrument, which contains the microphone and the receiver, is called handset.
- The handset is placed on the cradle when the telephone is not in use.
- In this position it opens the switches and disconnects the handset from the telephone system.
- An electromagnet, called the ringer is connected to the telephone line on the exchange side, so that a ring can be received from the exchange when it is called.
- The exchange determines that whether the telephone is idle or busy or initiating a call by monitoring the dc current.
- Microphone in telephone is regarded as transmitter. It is a transducer, which converts sound energy into electrical energy. There are different types of transmitter is the most widely used in the handset.
- The sound reproducer in telephone is called receiver. The receiver does the reverse function of the transmitter. It is a device, which converts electrical energy into sound energy.

FUNCTION OF SWITCHING SYSTEM-

- A switching system can be understand as a collection of switching elements arranged and controlled in such a way as to set up a common path between any two distant points.
- Switching systems reduced the complexity of wiring and made the telephony hassle free.

- The network connection cannot be simply made with telephone sets and bunch of wires, but a good system is required to make or break a connection. This system is known as the switching system or the switching office or the exchange.
- The subscribers instead of getting connected directly to one another, are connected to a switching office and then to the required subscriber.
- With the introduction of switching systems, the need for traditional connections between the subscribers reduced. All the subscribers need to have a connection with the switching system, which makes or breaks any connection, requested by the calling subscriber.
- The switching system, which is also called the telephone exchange, takes care of establishing the calls. Hence the total number of such links is equal to the number of subscribers connected to the system.
- Signaling is required for the switching system to establish or release a connection. It should also enable the switching system to detect whether a called subscriber is busy or not.
- The functions performed by a switching system in establishing and releasing connections are known as control function.
- The early systems required manual operations to establish telephone calls. An operator used to receive a call from the calling subscriber and then connect the call to the called subscriber. Later on, the system was automated.

STEPS OF CALL PROCEDURE-

Step 1:- Calling station goes off hook.

Step 2:- After detecting a dc current flow on the loop, the switching machine returns an audible dial tone to the calling station, acknowledging that the caller has access to the switching machine.

Step 3:- The caller dials the destination telephone number using one of two methods: mechanical dial pulsing or, more likely, electronic dual-tone multi frequency (Touch-Tone) signals.

Step 4:-When the switching machine detects the first dialed number, it removes the dial tone from the loop.

Step 5:-The switch interprets the telephone number and then locates the local loop for the destination telephone number.

Step 6:- Before ringing the destination telephone, the switching machine tests the destination loop for dc current to see if it is idle (on hook) or in use (off hook). At the same time, the switching machine locates a signal path through the switch between the two local loops.

Step 7a:- if the destination telephone is off hook, the switching machine sends a station busy signal back to the calling station.

Step7b:- If the destination telephone is on hook, the switching machine sends a ringing signal to the destination telephone on the local loop and at the same time sends a ring-back signal to the calling station to give the caller some assurance that something is happening.

Step 8:- When the destination answers the telephone, it completes the loop, causing dc current to flow.

Step 9:- The switch recognizes the dc current as the station answering the telephone. At this time, the switch removes the ringing and ring-back signals and completes the path through the switch, allowing the calling and called parties begin their conversation.

Step 10:- When either end goes on hook, the switching machine detects an open circuit on that loop and then drops the connections through the switch.

PRINCIPLE OF SPACE AND TIME SWITCHING-

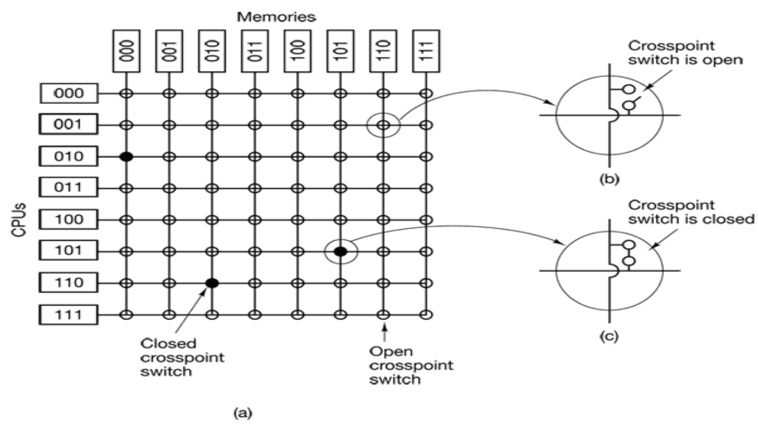
The switching scheme used by the electronic switching systems may be either **Space Division Switching** or **Time Division Switching.** In space division switching, a dedicated path is established between the calling and the called subscribers for the entire duration of the call. In time division switching, sampled values of speech signals are transferred at fixed intervals.

The time division switching may be analog or digital. In analog switching, the sampled voltage levels are transmitted as they are whereas in binary switching, they are binary coded and transmitted. If the coded values are transferred during the same time interval from input to output, the technique is called **Space Switching.** If the values are stored and transferred to the output at a late time interval, the technique is called as **Time Switching**. A time division digital switch may also be designed by using a combination of space and time switching techniques.

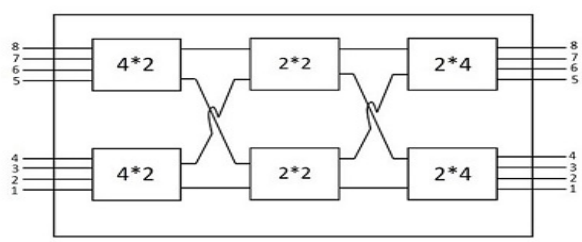
SPACE DIVISION SWITCHING

- In space division switching, the paths in the circuit are separated with each other spatially, i.e. different ongoing connections at a same instant of time, uses different switching paths.
- This was originally developed for the analog environment and has been carried over to the digital domain. The space switches are crossbar switches and multi stage switches.
- **Crossbar switch-**
  1. Basic building block of the switch is a metallic cross points or semiconductor gate that can be enabled or disabled by a control unit.
  2. The number of cross points grows with the square of the number of attached stations.
  3. Costly for a large switch.

4. The failure of a cross point prevents connection between the two devices whose lines intersect at that cross point.
5. The cross points are inefficiently utilized.
6. Only a small fraction of cross points are engaged even if all of the attached devices are active.



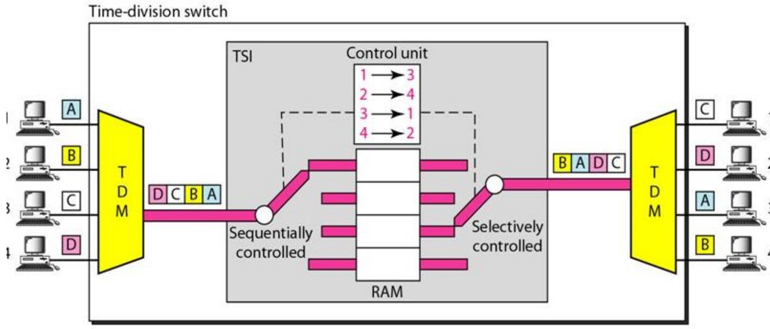
- Multistage space division switch-
  1. Some of the problem in crossbar switch can be overcome with the help of multistage space division switches.
  2. By splitting the crossbar switch into smaller units and interconnecting them it is possible to build multistage switches with fewer cross points.
  3. There is more than one path through the network to connect two endpoints, thereby increasing reliability.
  4. Multistage switches may lead to blocking.
  5. The problem may be tackled by increasing the number or size of the intermediate switches, which also increases the cost.



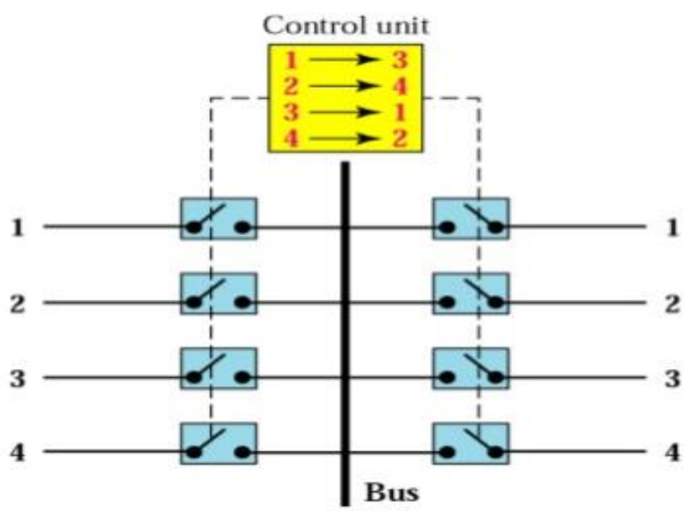
TIME DIVISION SWITCHING

- Both voice and data can be transmitted using digital signals.
- All modern circuit switches use digital time division multiplexing technique for establishing and maintaining circuits.
- Time division switching uses TDM to achieve switching i.e. different ongoing connections can use same switching path.
- There are two popular methods of time division switching i.e. **Time Slot Interchange (TSI)** and the **TDM Bus** .

- **TSI** changes the ordering of the slots based on desired connection and it has a random access memory to store data and flip the time slots.



- In **TDM bus** there are several input and outputs connected to a high speed bus.
- During a time slot only one particular output switch is closed, so only one connection at a particular instant of time.



<u>NUMBERING PLAN OF TELEPHONE NETWORKS (NATIONAL AND INTERNATIONAL NUMBERING)-</u>

Telephone Numbering Plan is a type of numbering scheme used in telecommunication to assign Phone Numbers to subscriber telephones or other telephony endpoints.

Telephone Numbering Plan defines the pattern of digits for a country phone number. Telephone Numbering Plan defines the **specific** components of phone numbers, display format styles, and codes.

Phone Numbers are the addresses of participants in a Telecommunications network, reachable by a system of destination code routing.

Telephone Numbering Plans are defined in each of administrative regions of the Public Switched Telephone Network (PSTN) and they are also present in private telephone networks. For public number systems, geographic location plays a role in the sequence of numbers assigned to each telephone subscriber.

Telephone Numbering Plan Components

Telephone Numbering Plan generally consists of the following Components: The E.164 recommendation provides the telephone number structure and functionality for three categories of telephone numbers used in international public telecommunication:

- Country Calling Code
- National Destination Code (Area Code)
- Subscriber Number

Country Calling Code-

Country Calling Code are a component of the international Telephone Numbering Plan and are prefixes for the member countries or regions of the International Telecommunication Union (ITU) and are defined by the ITU-T in standards E.123 and E.164. Country Calling Code are typically are necessary only when dialing a Phone Number to establish a call to another country and are dialed before the National Significant Number.

Country Calling Code prefixes enable International Direct Dialing (IDD), and are also referred to as International Subscriber Dialing (ISD) codes.

Country Calling Code, by convention, International Telephone Numbers are represented by prefixing the Country Calling Code with a plus sign (+), which also indicates to the subscriber that the local international call prefix must first be dialed.

Ordered by code

Each Country Calling Code's first digit directs to these broad areas:

- +1: North American Numbering Plan
- +2: mostly Africa
- +3–4: Europe
- +5: Americas outside the NANP
- +6: Southeast Asia and Oceania
- +7: Parts of the former Soviet Union
- +8: East Asia and special services
- +9: mostly southern Asia

Example-

Country Calling Code prefix in all countries belonging to the North American Numbering Plan is 011, while it is 00 in most European, Asian and African countries. On GSM (Mobile Networks), the prefix may automatically be inserted when the user prefixes a dialed number with the plus sign.

National Destination Code-

National Destination Code (NDC or NXX), commonly referred to as an International City Code or Number Plan Area or Area Code or significant leading digits of National Significant Number, is specified in ITU-T E.164.
National Destination Code identifies the Number Plan Area that is to be used.

National Destination Code optional code field which determined by each specific Telephone Numbering Plan.

National Destination Code when combined with the Subscriber Number (SN) - will constitute the National Significant Number within the international E.164-number for geographic areas.

National Destination Code North American Numbering Plan (NANP)

North American Numbering Plan (NANP) National Destination Code values are available.[2]
The following 555-XXXX National Destination Code line numbers remain in use:

- 555-1212 Directory Assistance National use
- 555-4334 Assigned National use
- 555-0100 through 555-0199 are **fictitious non-working numbers** reserved for entertainment/advertising.

Subscriber Number-

Subscriber Number is typically End-User unique Phone Number which may be in many different formats.
Subscriber Number is defined in the ITU E.164

PRIVATE BRANCH EXCHANGE (PBX)-

- PBX is a privately owned telephone switching system for handling multiple telephone lines without having to pay the phone company to lease each line separately.
- Normally a telephone line is connected to the phone company's local central office through "a trunk". The central office is responsible for routing incoming and outgoing calls. It provides other services like voice mail, call forwarding, caller ID and other features.
- Companies use a PBX for connecting all their internal phones to an external line. This way they can lease only one line and have many people using it, with each one having a phone at the desk with different number.
- The number is not in the same format as a phone number though, as it depends on the internal numbering. Inside a PBX, we only need to dial three digit or four digit numbers to make a call to another phone in the network.

Types of PBX-

TRADITIONAL PBX-

- A PBX is a privately owned phone system for handling multiple phone lines and routing calls.

- Depending on the services we need, it can be costly or rather cheap.
- These systems are being used less because they are harder to install and need to be connected to a land line.

HOSTED/ VIRTUAL PBX PHONE SYSTEM-

A hosted PBX is a virtual telephone system where, instead of having all our telephone hardware in the office, hardware is provided by a system hosting company and connect to the system via a network connection.

IP PBX-

- In this method costs are greatly reduced since there is no extra hardware to install.
- Phones can be connected to a computer port, rather than deals with wires.
- Like virtual PBX, we can manage the system online.
- Just install software and we are ready to start taking calls.
- Computer software allows for more features to be added.
- Low costs means its affordable for small and large companies.

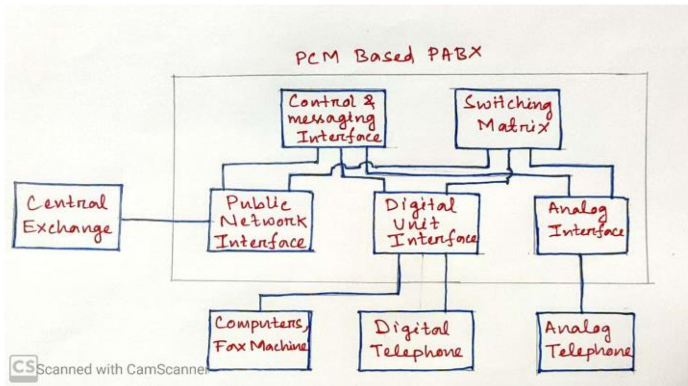
The equipment needed depends on the complexity and use of the PBX, for example, the types of phones used at a particular site. In general:

- Telephone trunk (multiple phone) lines that terminate at the PBX
- A computer with memory that manages the switching of the calls within the PBX and in and out of it
- The network of lines within the PBX
- Unified communications router (wireless and wired)
- Phone handset (USB, VoIP and SIP)
- VoIP gateway
- IP PBX
- Internet router
- Cables, cabinets, uninterruptible power supply
- Telephony application server

In some situations, alternatives to a PBX include Centrex service (in which a pool of lines are rented at the phone company's central office), key telephone systems, and, for very small enterprises, primary rate Integrated Services Digital Network (ISDN).

ELECTRONIC PRIVATE AUTOMATIC BRANCH EXCHANGE (EPABX)-

- EPABX provides telephone switching in an office or hospital.
- A PCM Based EPABX provides all the facilities like a Public Telephone Exchange.



Public Network Interface-

This section connects the EPABX with the local telephone Exchange and the no. of extensions connected with each line depends upon the design of the EPABX.

Control Message Interface-

This section checks the status of the subscriber's line and provides different tone.

Switching Matrix-

The function of switching matrix is to interconnect any two lines along with any internal signaling.

Digital Unit Interface-

The digital telephone sets, fax machines and other data base equipment are connected to this section.

Analog Unit Interface-

The analog telephone lines are connected with this section. The digital to analog and analog to digital conversion is done with the help of codes. The no. of extension point depends on the design of EPABX.

UNITS OF POWER MEASUREMENT

The **decibel** (symbol: **dB**) is a relative unit of measurement corresponding to one tenth of bel. It is used to express the ratio of one value of a power or field quantity to another, on a logarithmic scale, the logarithmic quantity being called the power level or field level, respectively.

INTERNET PROTOCOL TELEPHONY-

- Internet Protocol Telephony (IP Telephony) is the use of IP-based networks to build, provide and access voice, data or other forms of telephonic communications. IP telephony provides traditional telephonic communication over an IP-based network, the Internet - via an Internet service provider (ISP) - or directly from a telecommunications service provider.
- IP telephony is designed to replace the telecommunications' infrastructure of circuit switched public data networks (CSPDN) and public switched telephone networks (PSTN) with packet switched IP communication networks.
- In a consumer IP telephony solution, a soft IP phone application and backend Internet connection enable voice and data communication, such as calling and faxing. A user may call other softphone users, send or receive faxes and even communicate with circuit switched and cellular communication services.
- In an enterprise environment, IP telephony is implemented through physical IP phones that work on top of an IP network infrastructure. An IP phone's built-in firmware provides the complete functionality for initiating and managing telephonic communications. Moreover, IP telephony also supports video communication between two or more users.

Voice over Internet Protocol (VoIP), a popular IP telephony implementation, only supports voice communication over IP.

Advantages-

- IP telephony systems tend to be much less expensive than traditional systems for both local and long distance calling.
- Benefits of IP telephony systems to small businesses and organizations is that fact that they can be very flexible and have the ability to expand or contract with organization. If the business is growing rapidly or if the business is seasonal in nature.

- Internal communication is vital to organizations of every size and IP telephony systems make internal communication easier.

INTERNET TELEPHONY-

Internet Telephony is divided into three main categories consisting of –

(1) PC to PC telephony/ Calling

(2) PC to Phone telephony / Calling

(3) Phone to Phone telephony / Calling

PC to PC telephony / Calling-

- PC to PC telephony enables us to call another person who is online at the same time we are using the same telephony client.
- To use PC to PC calling, we will need compatible software, a microphone and speakers and internet access.
- PC to PC telephony is free. This makes it a popular choice for family and friends who live outside of each other's local telephone calling access.
- The client software converts transmitted speech into data packets and routes it over the internet. The receiving client turns the data packets back into voice signal. Example- Skype

PC to Phone Calling-

- PC to Phone telephony allows you to make calls from your computer to regular telephones. The technology is similar to PC to PC Calling.
- While this is a service we pay for, its usually cheaper than using a long distance telephone provider.
- To use PC to Phone services, we need to have the same equipment used for PC to PC calling, a microphone, speakers and internet access.
- The PC phone user's PC acts as the originating gateway, which converts the voice transmission into data packets onto the internet.
- At the end gateway provided by the software distributor, the data packets are converted to voice signals and routed to the Public Switched Telephone Network (PSTN)

Phone to Phone Calling-

- A relatively new type of Internet Telephony service, which is quickly gaining popularity.
- Phone to Phone telephony allow telephone calls to be placed over the internet, but if differs from the other types of internet telephony.
- Phone to Phone telephony does not require users to have special software, or even a computer to use it.
- Phone to Phone uses traditional telephones on both ends.
- Users with Phone to Phone calling place a call from their landline phone. The voice signals are digitized, compressed and converted to data packets.

- When the data packets arrive at the gateway, they are converted back into voice signals.
- Once converted, the voice signals are then transmitted through the local PSTN to the receiver.

UNIT- 5: DATA COMMUNICATION

Data communication refers to the exchange of data between a source and a receiver via form of transmission media such as a wire cable. Data communication is said to be local if communicating devices are in the same building or a similarly restricted geographical area.

The meanings of source and receiver are very simple. The device that transmits the data is known as source and the device that receives the transmitted data is known as receiver. Data communication aims at the transfer of data and maintenance of the data during the process but not the actual generation of the information at the source and receiver.

Components of data communication system-

A Communication system has following components:

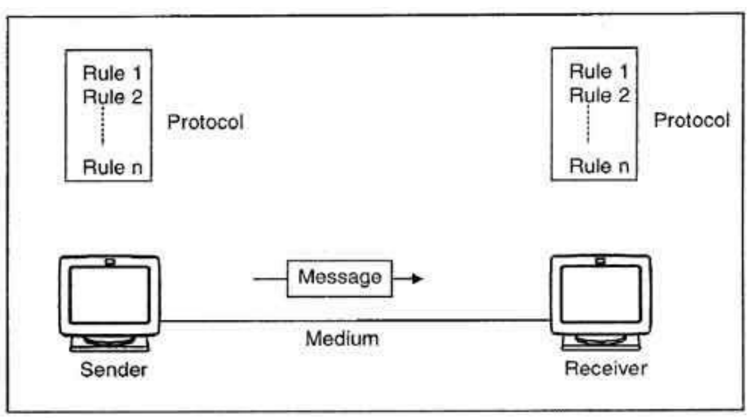
1. **Message**: It is the information or data to be communicated. It can consist of text, numbers, pictures, sound or video or any combination of these.

2. **Sender**: It is the device/computer that generates and sends that message.

3. **Receiver**: It 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 is carried from sender to the receiver. The medium can be wired like 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 govern the communication 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 duplication of packets, and to correctly identify 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 are 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 is discarded by the receiver and is retransmitted 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 is also built into most communication software packages. It prevents access of data by unauthorized users.

9. Log information. Several communication software are designed to develop log information, which consists of all jobs and data communications tasks that have taken place. Such information may be used for charging the users of the network based on their usage of the network resources.

The effectiveness depends on four fundamental characteristics of data communications

1. Delivery:
The system must deliver data to the exact destination. Data must not be received by other devices than the target device.

2. Accuracy:
The system must deliver data to the destination in a way that the target device receives the data accurately. If the protocol needs to alter the while in transmission, it must alter it back to its original form before representing it to the target device. The accuracy must be maintained.

3. Timeliness:
The system must deliver data in timely manner. Data delivered late can become useless. Data must be delivered as they are produced, in the order they are produced and without any significant delay.

4. Jitter:
Jitter refers to the variation of packet arrival time. Data is sent as packets, that is, a fixed amount of the whole data is sent in each to turn. These packets get joined back in the target device to

represent the complete data as it is. Each packet is sent with a predefined delay or acceptable amount delay. If packets are sent without maintaining the predefined delay then an uneven quality in the data might result.

PROTOCOL-

A protocol is basically a synonym for the rule. In Computer Networks, basically, communication occurs between entities in different systems. An entity is anything that is capable of sending or receiving information. Any two entities cannot simply send bit streams to each other and expect to be understood.

The entities must need to agree on a protocol in order of occurrence of the communication.

A Protocol is a set of rules that mainly govern data communications. The protocol mainly defines what is communicated, how it is communicated, and when it is communicated.

Key elements of a Protocol

The key elements of a protocol are as given below:

- **Syntax: –** Syntax refers to the structure or format of the data, meaning the order in which they are presented.
- **Semantics: –** Semantics refer to the meaning of each section of bits. How is a particular pattern to be interpreted, and what action is to be taken based on that interpretation?
- **Timing: –** Timing refers to two characteristics: when data should be sent and how fast they can be sent.

STANDARDS

Standards are essential in creating and maintaining an open and creative market for the equipment manufacturers and they also guarantee the national and international interoperability of the data, telecommunication technology, and process.

Standards are mainly used to provide guidelines to manufacturers, vendors, government agencies, and also to other service providers in order to ensure the kind of interconnectivity that is necessary for today's marketplace and also in international communications.

Data communication standards mainly fall into two categories:-

- **de facto** (which means "by fact"/ "by convention") Those standards have not been approved by an organized body but have been adopted as standards through widespread use. These types of standards are often established originally by the manufacturers who just seek to define the functionality of their new product or technology.
- **de jure** (which means "by law" /"by regulation") de jure standards are those standards that have been legislated by an officially recognized body.

Standard Organizations

Standards are mainly developed through the cooperation of Standard creation committees, government regulatory agencies, and forums.

Some Standard Creation committees are:

- International Organization of Standardization(ISO)
- American National Standards Institute(ANSI)
- Electronic Industries Association(EIA)
- Institute of Electrical and Electronics Engineers(IEEE)

STANDARDIZED PROTOCOL ARCHITECTURES

- Two standards:–
  - ✓ OSI Reference model
  - ✓ TCP/IP protocol suite

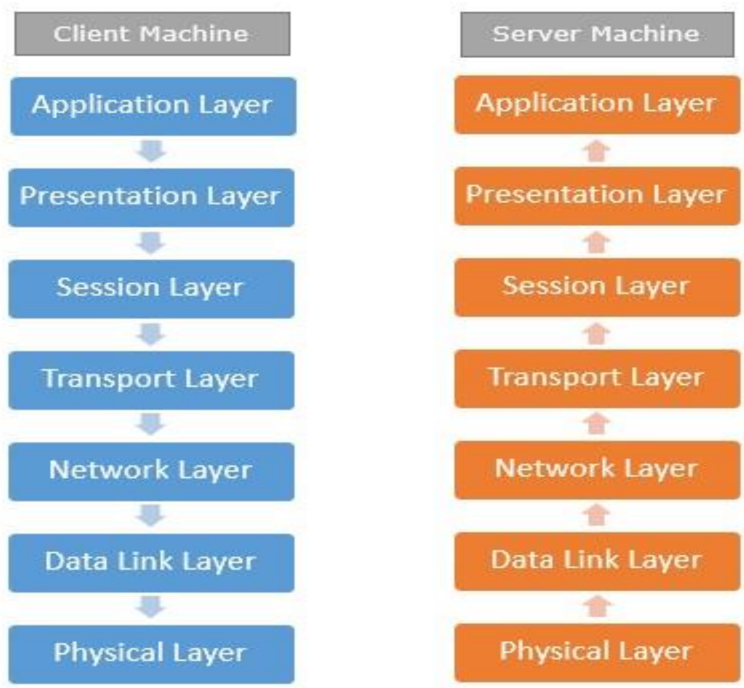
OSI Model

OSI model is not a **network architecture** because it does not specify the exact services and protocols for each layer. It simply tells what each layer should do by defining its input and output data. It is up to network architects to implement the layers according to their needs and resources available.

These are the seven layers of the OSI model –

- **Physical layer** –It is the first layer that physically connects the two systems that need to communicate. It transmits data in bits and manages simplex or duplex transmission by modem. It also manages Network Interface Card's hardware interface to the network, like cabling, cable terminators, topography, voltage levels, etc.
- **Data link layer** – It is the firmware layer of Network Interface Card. It assembles datagrams into frames and adds start and stop flags to each frame. It also resolves problems caused by damaged, lost or duplicate frames.
- **Network layer** – It is concerned with routing, switching and controlling flow of information between the workstations. It also breaks down transport layer datagrams into smaller datagrams.
- **Transport layer** – Till the session layer, file is in its own form. Transport layer breaks it down into data frames, provides error checking at network segment level and prevents a fast host from overrunning a slower one. Transport layer isolates the upper layers from network hardware.
- **Session layer** – This layer is responsible for establishing a session between two workstations that want to exchange data.

- **Presentation layer –** This layer is concerned with correct representation of data, i.e. syntax and semantics of information. It controls file level security and is also responsible for converting data to network standards.
- **Application layer –** It is the topmost layer of the network that is responsible for sending application requests by the user to the lower levels. Typical applications include file transfer, E-mail, remote logon, data entry, etc.



It is not necessary for every network to have all the layers. For example, network layer is not there in broadcast networks.

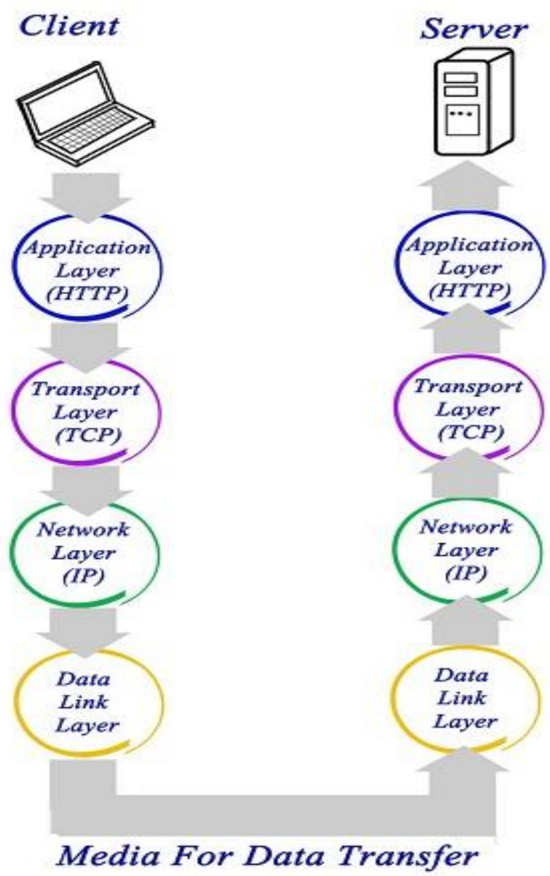
When a system wants to share data with another workstation or send a request over the network, it is received by the application layer. Data then proceeds to lower layers after processing till it reaches the physical layer.

At the physical layer, the data is actually transferred and received by the physical layer of the destination workstation. There, the data proceeds to upper layers after processing till it reaches application layer.

At the application layer, data or request is shared with the workstation. So each layer has opposite functions for source and destination workstations. For example, data link layer of the source workstation adds start and stop flags to the frames but the same layer of the destination workstation will remove the start and stop flags from the frames.

TCP/IP

TCP/IP stands for **Transmission Control Protocol/Internet Protocol**. TCP/IP is a set of layered protocols used for communication over the Internet. The communication model of this suite is client-server model. A computer that sends a request is the client and a computer to which the request is sent is the server.



TCP/IP has four layers –

- **Application layer –** Application layer protocols like HTTP and FTP are used.
- **Transport layer –** Data is transmitted in form of datagrams using the Transmission Control Protocol (TCP). TCP is responsible for breaking up data at the client side and then reassembling it on the server side.
- **Network layer –** Network layer connection is established using Internet Protocol (IP) at the network layer. Every machine connected to the Internet is assigned an address called IP address by the protocol to easily identify source and destination machines.
- **Data link layer –** Actual data transmission in bits occurs at the data link layer using the destination address provided by network layer.

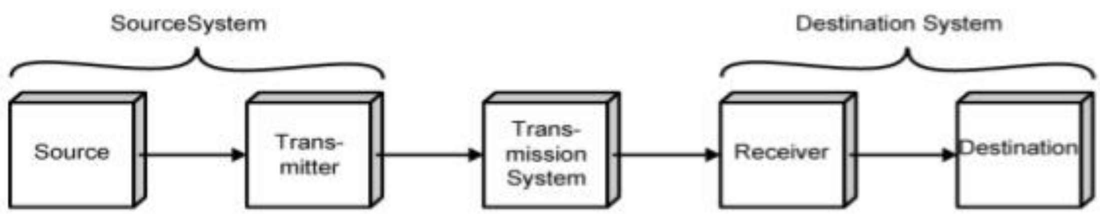
TCP/IP is widely used in many communication networks other than the Internet.

DATA COMMUNICATION CIRCUITS

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

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

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



(a) General block diagram

(b) Example

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

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

Receiver: - The receiver accepts the signal from the transmission medium and converts it into a form that can be handled by the destination device. For example, a modem will 10 Data C Unit-1

Introduction to Data Communications accept an analog signal coming from a network or transmission line and convert it into a digital bit stream.

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

TYPES OF TRANSMISSION & TRANSMISSION MODE-

Data Transmission mode defines the direction of the flow of information between two communication devices. It is also called Data Communication or Directional Mode. It specifies the direction of the flow of information from one place to another in a computer network.

In the Open System Interconnection (OSI) Layer Model, the **Physical Layer** is dedicated to data transmission in the network. It mainly decides the direction of data in which the data needs to travel to reach the receiver system or node.

So, in this blog, we will learn about different data transmission modes based on the direction of exchange, synchronization between the transmitter and receiver, and the number of bits sent simultaneously in a computer network.

The data transmission modes can be characterized in the following three types based on the direction of exchange of information:

1. Simplex
2. Half-Duplex
3. Full Duplex

The data transmission modes can be characterized in the following two types based on the synchronization between the transmitter and the receiver:

1. Synchronous
2. Asynchronous

The data transmission modes can be characterized in the following two types based on the number of bits sent simultaneously in the network:

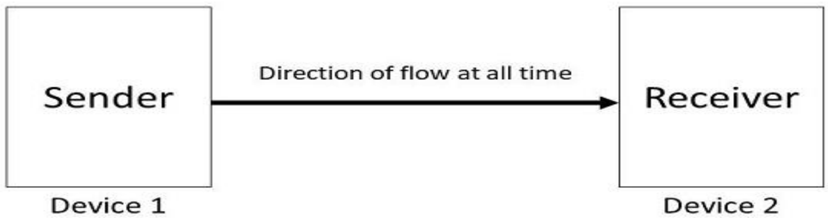
1. Serial
2. Parallel

Now, let us study these various data transmission modes in the computer network one by one.

According to the Direction of Exchange of Information:

1. Simplex

Simplex is the data transmission mode in which the data can flow only in one direction, i.e., the communication is unidirectional. In this mode, a sender can only send data but cannot receive it. Similarly, a receiver can only receive data but cannot send it.



Simplex Mode

This transmission mode is not so popular because we cannot perform two-way communication between the sender and receiver in this mode. It is mainly used in the business field as in sales that do not require any corresponding reply. It is similar to a one-way street.

For Example, Radio and TV transmission, keyboard, mouse, etc.

Following are the advantages of using a Simplex transmission mode:

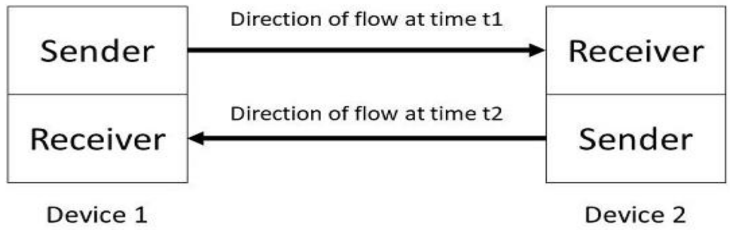
1. It utilizes the full capacity of the communication channel during data transmission.
2. It has the least or no data traffic issues as data flows only in one direction.

Following are the disadvantages of using a Simplex transmission mode:

1. It is unidirectional in nature having no inter-communication between devices.
2. There is no mechanism for information to be transmitted back to the sender(No mechanism for acknowledgement).

2. Half-Duplex

Half-Duplex is the data transmission mode in which the data can flow in both directions but in one direction at a time. It is also referred to as Semi-Duplex. In other words, each station can both transmit and receive the data but not at the same time. When one device is sending the other can only receive and vice-versa.



Half-Duplex Mode

In this type of transmission mode, the entire capacity of the channel can be utilized for each direction. Transmission lines can carry data in both directions, but the data can be sent only in one direction at a time.

This type of data transmission mode can be used in cases where there is no need for communication in both directions at the same time. It can be used for error detection when the sender does not send or the receiver does not receive the data properly. In such cases, the data needs to be transmitted again by the receiver.

For Example, Walkie-Talkie, Internet Browsers, etc.

Following are the advantages of using a half-duplex transmission mode:

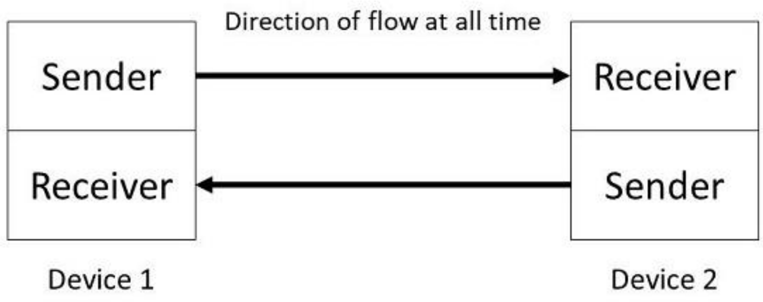
1. It facilitates the optimum use of the communication channel.
2. It provides two-way communication.

Following are the disadvantages of using a half-duplex transmission mode:

1. The two-way communication cannot be established simultaneously at the same time.
2. Delay in transmission may occur as only one way communication can be possible at a time.

3. Full-Duplex

Full-Duplex is the data transmission mode in which the data can flow in both directions at the same time. It is bi-directional in nature. It is two-way communication in which both the stations can transmit and receive the data simultaneously.



Full-Duplex Mode

Full-Duplex mode has double bandwidth as compared to the half-duplex. The capacity of the channel is divided between the two directions of communication. This mode is used when communication in both directions is required simultaneously.

For Example, a Telephone Network, in which both the persons can talk and listen to each other simultaneously.

Following are the advantages of using a full-duplex transmission mode:

1. The two-way communication can be carried out simultaneously in both directions.
2. It is the fastest mode of communication between devices.

Following are the disadvantages of using a half-duplex transmission mode:

1. The capacity of the communication channel is divided into two parts. Also, no dedicated path exists for data transfer.
2. It has improper channel bandwidth utilization as there exist two separate paths for two communicating devices.

According to the synchronization between the transmitter and the receiver:

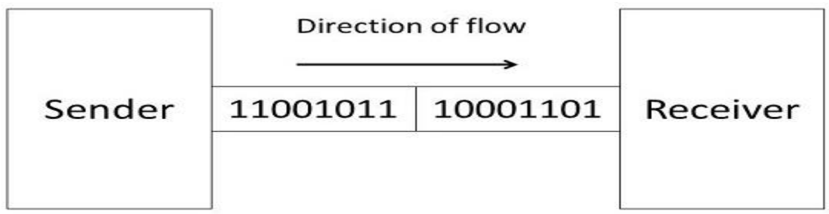
1. Synchronous

The Synchronous transmission mode is a mode of communication in which the bits are sent one after another without any start/stop bits or gaps between them. Actually, both the sender and receiver are paced by the same system clock. In this way, synchronization is achieved.

In a Synchronous mode of data transmission, bytes are transmitted as blocks in a continuous stream of bits. Since there is no start and stop bits in the message block. It is the responsibility of the receiver to group the bits correctly. The receiver counts the bits as they arrive and groups them in eight bits unit. The receiver continuously receives the information at the same rate that the transmitter has sent it. It also listens to the messages even if no bits are transmitted.

In synchronous mode, the bits are sent successively with no separation between each character, so it becomes necessary to insert some synchronization elements with the message, this is called "Character-Level Synchronization".

For Example, if there are two bytes of data, say(10001101, 11001011) then it will be transmitted in the synchronous mode as follows:



Synchronous Mode

For Example, communication in CPU, RAM, etc.

Following are the advantages of using a Synchronous transmission mode:

1. Transmission speed is fast as there is no gap between the data bits.

Following are the disadvantages of using a Synchronous transmission mode:

1. It is very expensive.

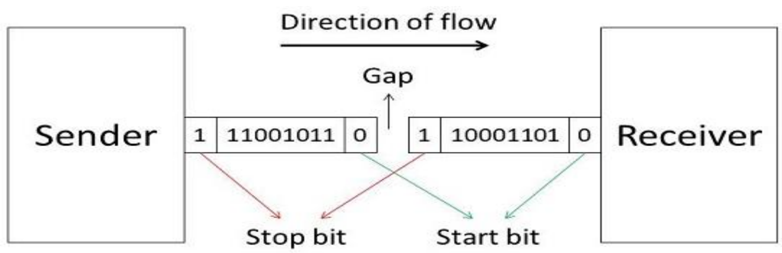
2. Asynchronous

The Asynchronous transmission mode is a mode of communication in which a start and the stop bit is introduced in the message during transmission. The start and stop bits ensure that the data is transmitted correctly from the sender to the receiver.

Generally, the start bit is '0' and the end bit is '1'.Asynchronous here means 'asynchronous at the byte level', but the bits are still synchronized. The time duration between each character is the same and synchronized.

In an asynchronous mode of communication, data bits can be sent at any point in time. The messages are sent at irregular intervals and only one data byte can be sent at a time. This type of transmission mode is best suited for short-distance data transfer.

For Example, if there are two bytes of data, say(10001101, 11001011) then it will be transmitted in the asynchronous mode as follows:



Asynchronous Mode

For Example, Data input from a keyboard to the computer.

Following are the advantages of using an Asynchronous transmission mode:

1. It is a cheap and effective mode of transmission.
2. Data transmission accuracy is high due to the presence of start and stop bits.

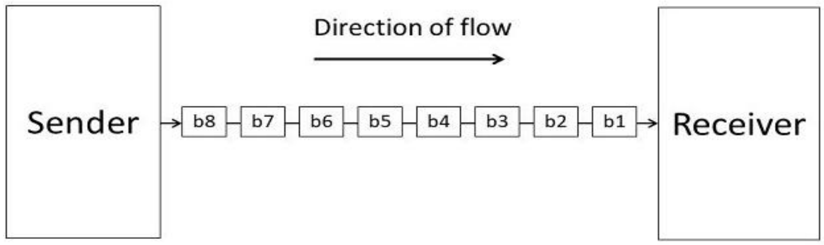
Following are the disadvantages of using an Asynchronous transmission mode:

1. The data transmission can be slower due to the gaps present between different blocks of data.

According to the number of bits sent simultaneously in the network:

1. Serial

The Serial data transmission mode is a mode in which the data bits are sent serially one after the other at a time over the transmission channel.



Serial Mode

It needs a single transmission line for communication. The data bits are received in synchronization with one another. So, there is a challenge of synchronizing the transmitter and receiver.

In serial data transmission, the system takes several clock cycles to transmit the data stream. In this mode, the data integrity is maintained, as it transmits the data bits in a specific order, one after the other.

This type of transmission mode is best suited for long-distance data transfer, or the amount of data being sent is relatively small.

For Example, Data transmission between two computers using serial ports.

Following are the advantages of using a serial transmission mode:

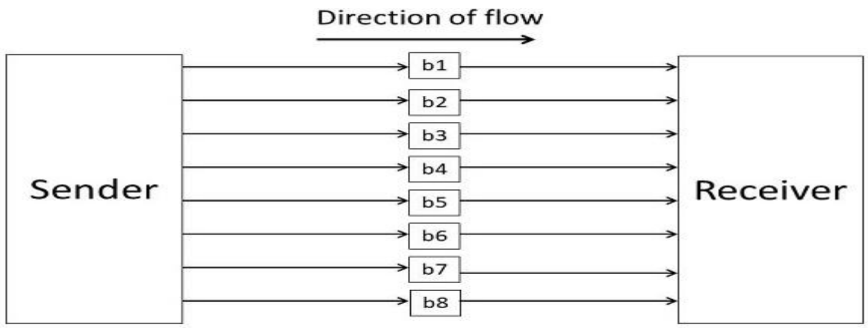
1. It can be used for long-distance data transmission as it is reliable.
2. The number of wires and complexity is less.
3. It is cost-effective.

Following are the disadvantages of using a serial transmission mode:

1. The Data transmission rate is slow due to a single transmission channel.

2. Parallel

The Parallel data transmission mode is a mode in which the data bits are sent parallelly at a time. In other words, there is a transmission of n-bits at the same time simultaneously.



Parallel Mode

Multiple transmission lines are used in such modes of transmission. So, multiple data bytes can be transmitted in a single system clock. This mode of transmission is used when a large amount of data has to be sent in a shorter duration of time. It is mostly used for short-distance communication.

For n-bits, we need n-transmission lines. So, the complexity of the network increases but the transmission speed is high. If two or more transmission lines are too close to each other, then there may be a chance of interference in the data, degrading the signal quality.

For Example, Data transmission between computer and printer.

Following are the advantages of using a parallel transmission mode:

1. It is easy to program or implement.
2. Data transmission speed is high due to the n-transmission channel.

Following are the disadvantages of using a parallel transmission mode:

1. It requires more transmission channels, and hence cost-ineffective.
2. Interference in data bits, likewise in video conferencing.

DATA COMMUNICATION CODES-

line coding schemes divided into five categories:

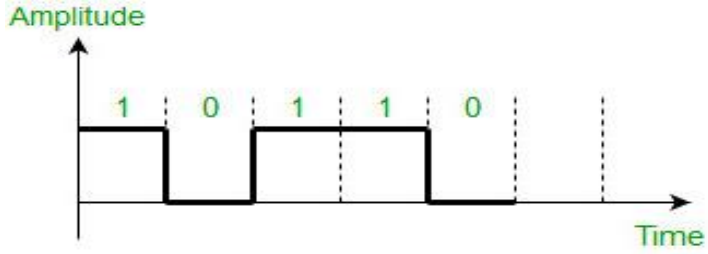
1. Unipolar (eg. NRZ scheme).
2. Polar (eg. NRZ-L, NRZ-I, RZ, and Biphase – Manchester and differential Manchester).
3. Bipolar (eg. AMI and Pseudo ternary).
4. Multilevel
5. Multi-transition

Unipolar scheme –

In this scheme, all the signal levels are either above or below the axis.

- **Non return to zero (NRZ) –**

  It is unipolar line coding scheme in which positive voltage defines bit 1 and the zero voltage defines bit 0. Signal does not return to zero at the middle of the bit thus it is called NRZ. For example: Data = 10110.



  But this scheme uses more power as compared to polar scheme to send one bit per unit line resistance. Moreover for continuous set of zeros or ones there will be self-synchronization and base line wandering problem.

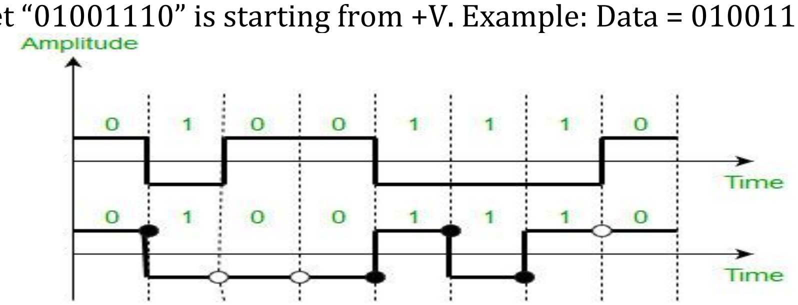
Polar schemes –

In polar schemes, the voltages are on the both sides of the axis.

- **NRZ-L and NRZ-I –**

  These are somewhat similar to unipolar NRZ scheme but here we use two levels of amplitude (voltages). For **NRZ-L(NRZ-Level)**, the level of the voltage determines the value of the bit, typically binary 1 maps to logic-level high, and binary 0 maps to logic-level low, and for **NRZ-I(NRZ-Invert)**, two-level signal has a transition at a boundary if the next bit that we are going to transmit is a logical 1, and does not have a transition if the next bit that we are going to transmit is a logical 0.

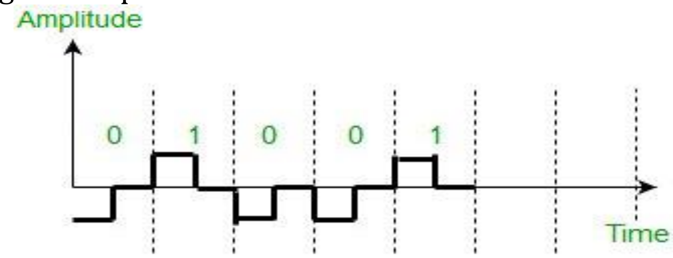
  Note – For NRZ-I we are assuming in the example that previous signal before starting of data set "01001110" was positive. Therefore, there is no transition at the beginning and first bit "0" in current data set "01001110" is starting from +V. Example: Data = 01001110.



  Comparison between NRZ-L and NRZ-I: Baseline wandering is a problem for both of them, but for NRZ-L it is twice as bad as compared to NRZ-I. This is because of transition at the boundary for NRZ-I (if the next bit that we are going to transmit is a logical 1). Similarly self-synchronization problem is similar in both for long sequence of 0's, but for long sequence of 1's it is more severe in NRZ-L.

- **Return to zero (RZ) –**

  One solution to NRZ problem is the RZ scheme, which uses three values positive, negative and zero. In this scheme signal goes to 0 in the middle of each bit. **Note –** The logic we are using here to represent data is that for bit 1 half of the signal is represented by +V and half by zero voltage and for bit 0 half of the signal is represented by -V and half by zero voltage. Example: Data = 01001.



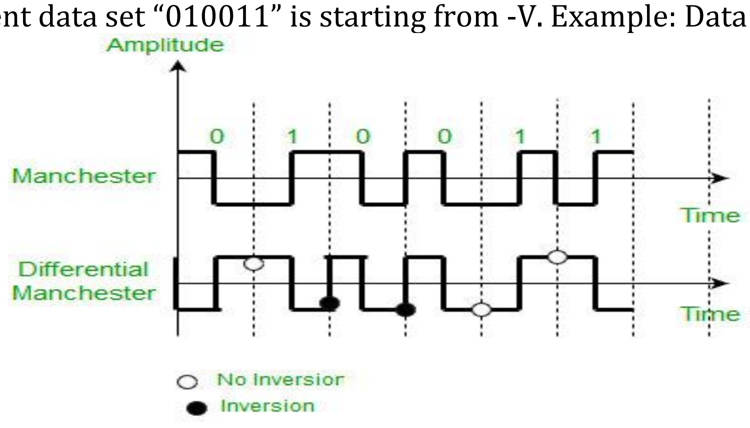
Main disadvantage of RZ encoding is that it requires greater bandwidth. Another problem is the complexity as it uses three levels of voltage. As a result of all these deficiencies, this scheme is not used today. Instead, it has been replaced by the better-performing Manchester and differential Manchester schemes.

- **Bi-phase (Manchester and Differential Manchester) –**

Manchester encoding is somewhat combination of the RZ (transition at the middle of the bit) and NRZ-L schemes. The duration of the bit is divided into two halves. The voltage remains at one level during the first half and moves to the other level in the second half. The transition at the middle of the bit provides synchronization.

Differential Manchester is somewhat combination of the RZ and NRZ-I schemes. There is always a transition at the middle of the bit but the bit values are determined at the beginning of the bit. If the next bit is 0, there is a transition, if the next bit is 1, there is no transition.

Note –
1. The logic we are using here to represent data using Manchester is that for bit 1 there is transition form -V to +V volts in the middle of the bit and for bit 0 there is transition from +V to -V volts in the middle of the bit.
2. For differential Manchester we are assuming in the example that previous signal before starting of data set "010011" was positive. Therefore there is transition at the beginning and first bit "0" in current data set "010011" is starting from -V. Example: Data = 010011.



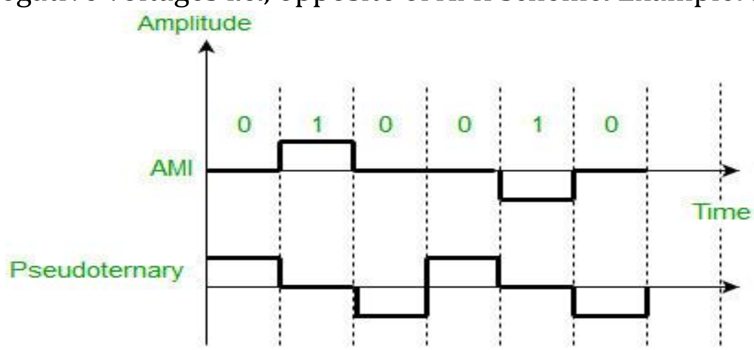
The Manchester scheme overcomes several problems associated with NRZ-L, and differential Manchester overcomes several problems associated with NRZ-I as there is no baseline wandering and no DC component because each bit has a positive and negative voltage contribution.

Only limitation is that the minimum bandwidth of Manchester and differential Manchester is twice that of NRZ.

Bipolar schemes –

In this scheme there are three voltage levels positive, negative, and zero. The voltage level for one data element is at zero, while the voltage level for the other element alternates between positive and negative.

- **Alternate Mark Inversion (AMI) –** A neutral zero voltage represents binary 0. Binary 1's are represented by alternating positive and negative voltages.
- **Pseudo ternary –** Bit 1 is encoded as a zero voltage and the bit 0 is encoded as alternating positive and negative voltages i.e., opposite of AMI scheme. Example: Data = 010010.



The bipolar scheme is an alternative to NRZ. This scheme has the same signal rate as NRZ, but there is no DC component as one bit is represented by voltage zero and other alternates every time.

Multilevel Scheme-

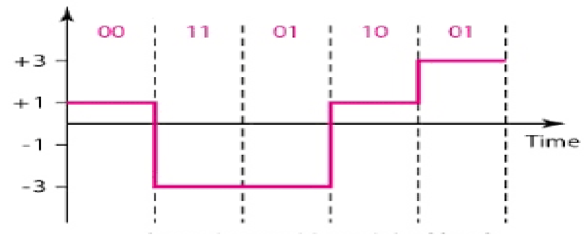
2BIQ-

In two binary one quaternary, uses data patterns of size two and encodes the two bit patterns as one signal element belonging to a four level signal.

	Previous level: positive	Previous level: negative
Next bits	Next level	Next level
00	+1	-1
01	+3	-3
10	-1	+1
11	-3	+3

Transition table

Ex.- Data- 0011011001



8B6T- (Eight binary, six ternary)

- This code is used with 100BASE-4T cable. The idea is to encode a pattern of 8 bits as a pattern of 8 bits as a pattern of six signal elements.
- Each signal pattern has a weight of 0 or +1.
- The three possible signal levels are represented as -, 0 and +.

4D-PAM5- (Four dimensional five level pulse amplitude modulation)

- The 4D means that data is sent over four wires at the same time.
- It uses five voltage levels, such as -2, -1, 0,1 and 2.
- However, one level, level 0 is used only for forward error detection.
- The technique is designed to send data over four channels.
- Gigabit LANs use this technique to send 1-Gbps data over four copper cables that can handle 125 Mband.

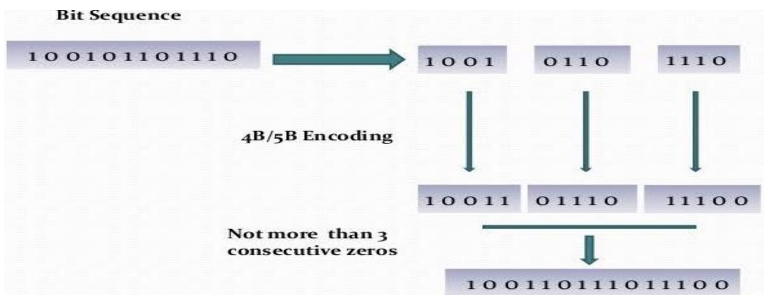
BLOCK CODING-

Block coding is normally referred to as mB/nB coding, it replaces each m bit group with an n bit group.

4B/5B (Four binary/ Five binary)-

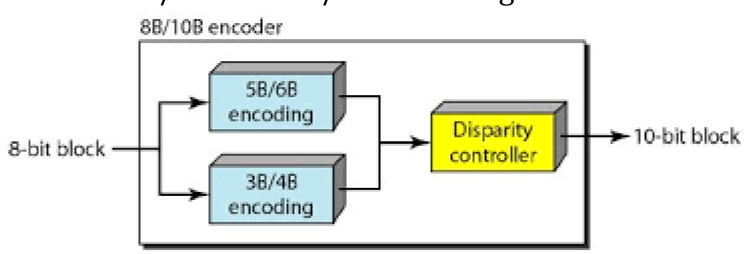
- In 4B/5B, the 5 bit output that replaces the 4 bit input has no more than one leading zero and no more than two trailing zeros.0
- So, when different groups are combined to make a new sequence, there are never more than three consecutive 0s.

Data Sequence	Encoded Sequence	Control Sequence	Encoded Sequence
0000	11110	Q (Quiet)	00000
0001	01001	I (Idle)	11111
0010	10100	H (Halt)	00100
0011	10101	J (Start delimiter)	11000
0100	01010	K (Start delimiter)	10001
0101	01011	T (End delimiter)	01101
0110	01110	S (Set)	11001
0111	01111	R (Reset)	00111
1000	10010		
1001	10011		
1010	10110		
1011	10111		
1100	11010		
1101	11011		
1110	11100		
1111	11101		



8B/10B (Eight binary/ ten binary)-

- This is similar to 4B/5B encoding except that a group of 8 bits of data is substituted by a 10 bit code.
- It provides greater error detection capability than 4B/ 5B. the 8B/ 10B block coding is actually a combination of 5B/ 6B and 3B/ 4B encoding.



SCRAMBLING-

Scrambling is a technique that does not increase the number of bits and does provide synchronization. Problem with technique like Bipolar AMI (Alternate Mark Inversion) is that continuous sequence of zero's create synchronization problems one solution to this is Scrambling.

There are two common scrambling techniques:

1. B8ZS(Bipolar with 8-zero substitution)
2. HDB3(High-density bipolar3-zero)

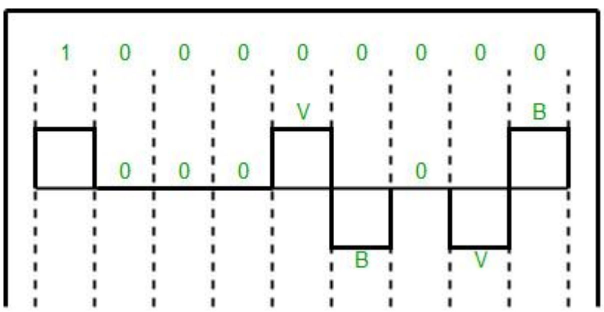
B8ZS (Bipolar with 8-zero substitution) –

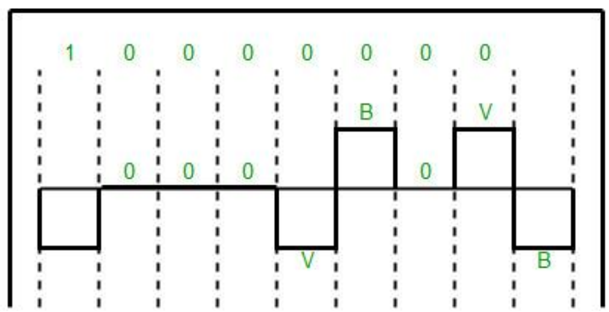
This technique is similar to Bipolar AMI except when eight consecutive zero-level voltages are encountered they are replaced by the sequence,"000VB0VB".

Note –

- V (Violation), is a non-zero voltage which means signal have same polarity as the previous non-zero voltage. Thus it is violation of general AMI technique.
- B (Bipolar), also non-zero voltage level which is in accordance with the AMI rule (i.e., opposite polarity from the previous non-zero voltage).

Example: Data = 100000000





Note – Both figures (left and right one) are correct, depending upon last non-zero voltage signal of previous data sequence (i.e., sequence before current data sequence "100000000").

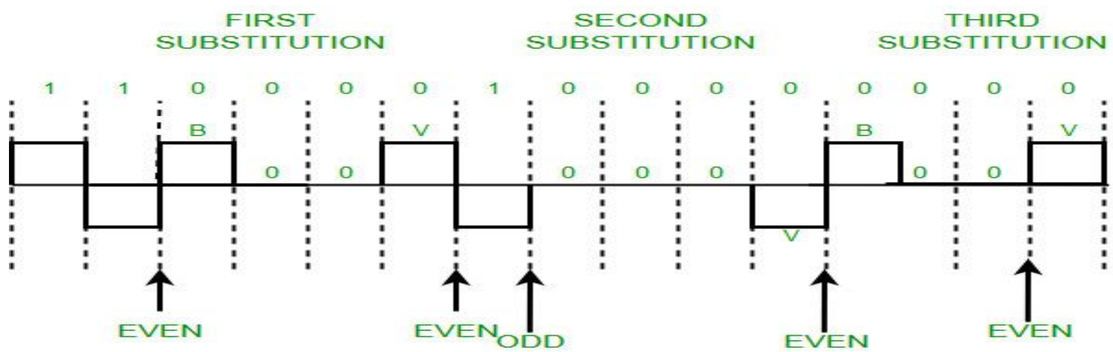
HDB3(High-density bipolar3-zero) –

In this technique four consecutive zero-level voltages are replaced with a sequence "000V" or "B00V".

Rules for using these sequences:

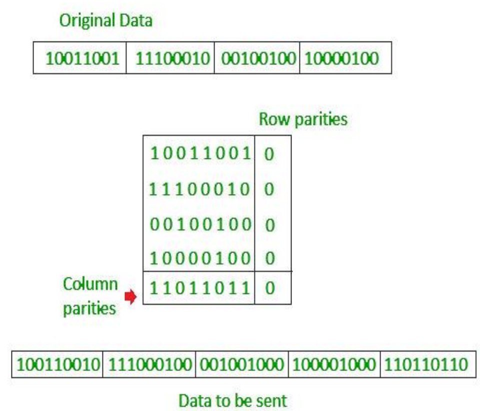
- If the number of nonzero pulses after the last substitution is odd, the substitution pattern will be "000V", this helps maintaining total number of nonzero pulses even.
- If the number of nonzero pulses after the last substitution is even, the substitution pattern will be "B00V". Hence even number of nonzero pulses is maintained again.

Example: Data = 1100001000000000



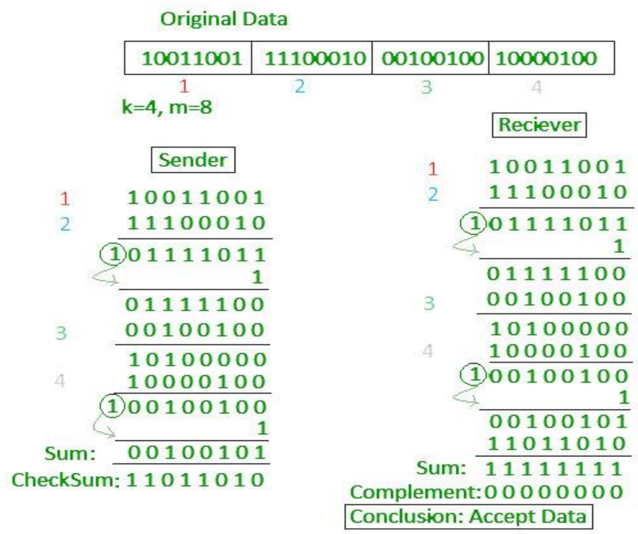
Explanation – After representing first two 1's of data we encounter four consecutive zeros. Since our last substitutions were two 1's (thus number of non-zero pulses is even).So, we substitute four zeros with "B00V".

BASIC IDEA OF ERROR CONTROL & ERROR DETECTION-



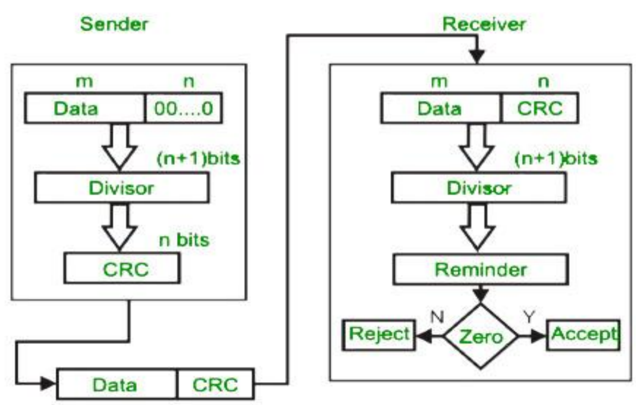
Checksum

- In checksum error detection scheme, the data is divided into k segments each of m bits.
- In the sender's end the segments are added using 1's complement arithmetic to get the sum. The sum is complemented to get the checksum.
- The checksum segment is sent along with the data segments.
- At the receiver's end, all received segments are added using 1's complement arithmetic to get the sum. The sum is complemented.
- If the result is zero, the received data is accepted; otherwise discarded.

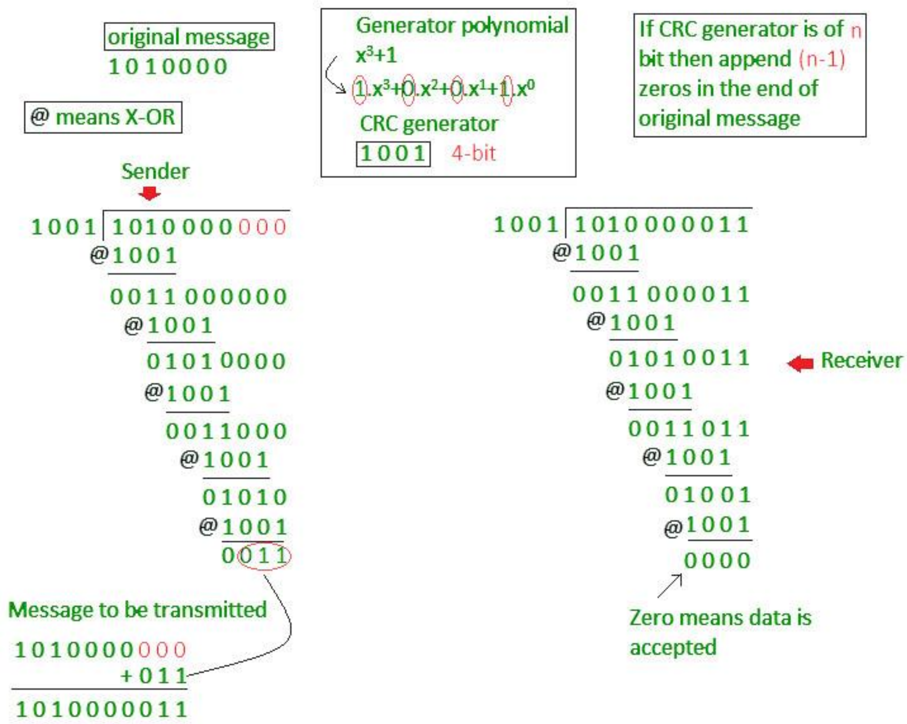


Cyclic redundancy check (CRC)

- Unlike checksum scheme, which is based on addition, CRC is based on binary division.
- In CRC, a sequence of redundant bits, called cyclic redundancy check bits, are appended to the end of data unit so that the resulting data unit becomes exactly divisible by a second, predetermined binary number.
- At the destination, the incoming data unit is divided by the same number. If at this step there is no remainder, the data unit is assumed to be correct and is therefore accepted.
- A remainder indicates that the data unit has been damaged in transit and therefore must be rejected.



Example:



ERROR CORRECTION

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 handled in two ways:

- **Backward error correction:** Once the error is discovered, 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:

$$2^r >= d+r+1$$

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.

Relationship b/w Error position & binary number.

Error Position	Binary Number
0	000
1	001
2	010
3	011
4	100
5	101
6	110
7	111

Let's 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: $2^r >= d+r+1$

$2^r >= 4+r+1$

Therefore, the value of r is 3 that satisfies the above relation.

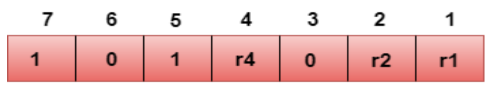
Total number of bits = d+r = 4+3 = 7;

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, 2^1, 2^2**.

1. The position of r1 = 1
2. The position of r2 = 2
3. The position of r4 = 4

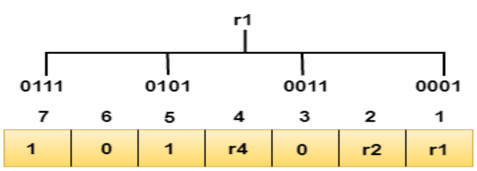
Representation of Data on the addition of parity bits:



Determining the Parity bits

Determining the r1 bit

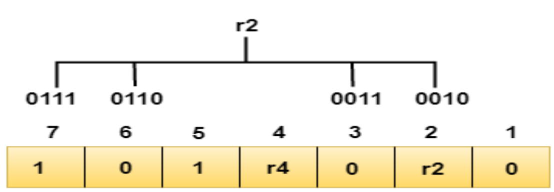
The r1 bit is calculated by performing a parity check on the bit positions whose binary representation includes 1 in the first position.



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

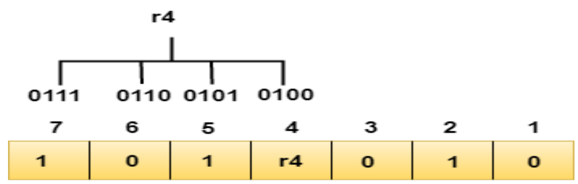
The r2 bit is calculated by performing a parity check on the bit positions whose binary representation includes 1 in the second position.



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**.

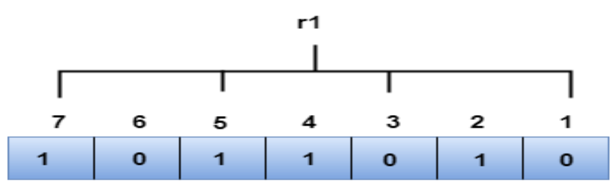
Data transferred is given below:

7	6	5	4	3	2	1
1	0	1	0	0	1	0

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

R1 bit

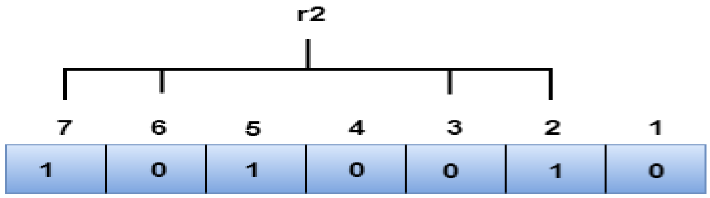
The bit positions of the r1 bit are 1,3,5,7



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

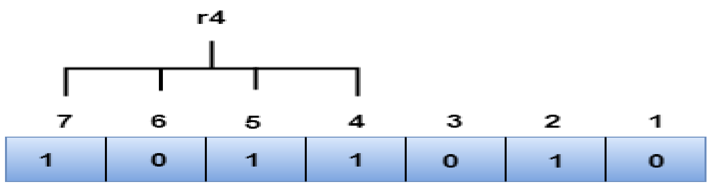
The bit positions of r2 bit are 2,3,6,7.



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.

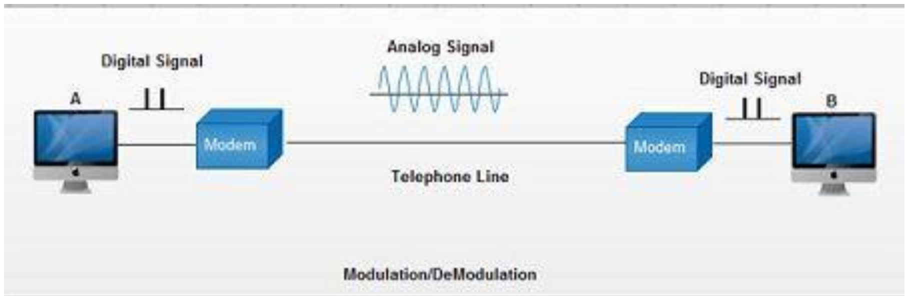
- The binary representation of redundant bits, i.e., r4r2r1 is 100, and its corresponding decimal value is 4. Therefore, the error occurs in a 4^{th} bit position. The bit value must be changed from 1 to 0 to correct the error.

MODEM & ITS BASIC BLOCK DIAGRAM

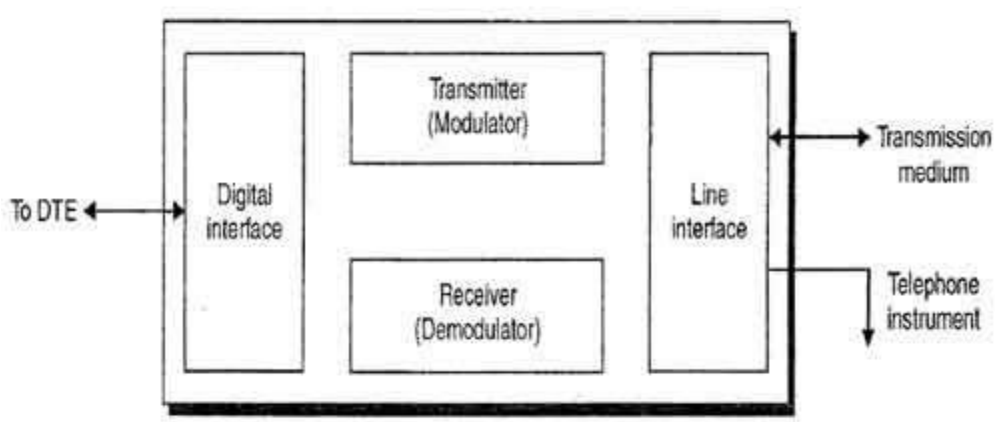
Modem is abbreviation for Modulator – De-modulator. Modems are used for data transfer from one computer network to another computer network through telephone lines. The computer network works in digital mode, while analog technology is used for carrying massages across phone lines.

Modulator converts information from digital mode to analog mode at the transmitting end and de-modulator converts the same from analog to digital at receiving end. The process of converting analog signals of one computer network into digital signals of another computer network so they can be processed by a receiving computer is referred to as digitizing.

When an analog facility is used for data communication between two digital devices called Data Terminal Equipment (DTE), modems are used at each end. DTE can be a terminal or a computer.



The modem at the transmitting end converts the digital signal generated by DTE into an analog signal by modulating a carrier. This modem at the receiving end demodulates the carrier and hand over the demodulated digital signal to the DTE.



Building blocks of a modem

The transmission medium between the two modems can be dedicated circuit or a switched telephone circuit. If a switched telephone circuit is used, then the modems are connected to the local telephone exchanges. Whenever data transmission is required connection between the modems is established through telephone exchanges.

VOICE BAND MODEM

Most modems are "voice band"; i.e., they enable digital terminal equipment to communicate over telephone channels, which are designed around the narrow bandwidth requirements of the human voice.

Voice band modems are marketed as freestanding, book-sized modules that plug into a telephone or cable outlet and a port on a personal computer.

Voice band modems are installed as circuit boards directly into computers and fax machines. They are also available as small card-sized units that plug into laptop computers.

Voice-band modems were introduced in the late 1950s to transmit data through the public switched telephone network (PSTN).

The first voice-band modem, used for interconnecting computers was the Bell 202, which provided a data rate of 1200 bps. This was a four-wire modem, where two two-wire pairs are used in simplex mode (i.e., transmission takes place one way permanently), one for each direction.

The recent V.34 modem provides a data rate of 33,600 bps over classical two-wire dial-up telephone lines. This is a two-wire full-duplex modem. By transmitting 33,600 bps in 3:6 kHz voice band, V.34 modems send nearly 10 bps/Hz.

Voice-band modem uses passband data transmission, where a baseband (band around zero frequency) information signal is converted, by modulation, into a passband (band away from zero frequency) signal which can transit a passband channel, and a demodulation is performed in the receiver to recover the baseband information signal.

Here the channel refers to physical communication medium, which is the telephone line in the case of voice-band modems.

UNIT- 6: WIRELESS COMMUNICATION

BASIC CONCEPT OF CELL PHONE-

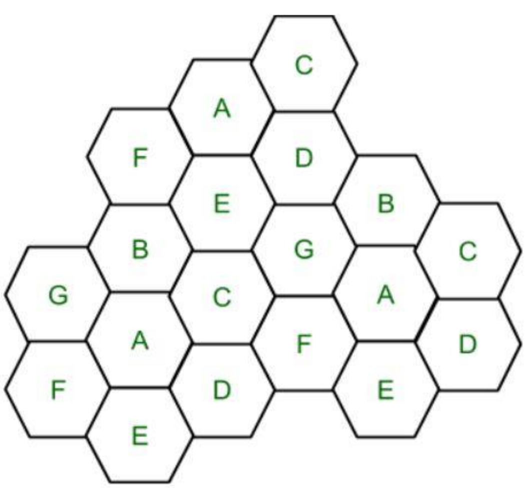
A cellular phone is a telecommunication device that uses radio waves over a networked area and is served through a base station at a fixed location, enabling calls to transmit wirelessly over a wide range to a fixed landline or via the internet.

In this networked system, the cellular phone is identified as a mobile system consisting of the equipment and SIM card that actually assigns the mobile telephone number. A cellular phone is also known as a cellphone or mobile phone.

Cellular network is technology for mobile phones, personal communication systems, wireless networking etc. cellular networks use lower power, shorter range and more transmitters for data transmission.

The coverage area of cellular networks are divided into cells, each cell having its own frequencies. Data communication in cellular networks is served by its base station transmitter, receiver and its control unit. A hexagon cell shape is highly recommended for its easy coverage and calculations.

FREQUENCY REUSE-



Cell with the same letter uses the same set of channels. Then Total number of channels (S) will be,

$S = kN$

k = Channels allocated to each cell ($k<S$)
N = Total number of cells or Cluster Size

In the above diagram cluster size is 7 (A,B,C,D,E,F,G) thus frequency reuse factor is 1/7.

N is the number of cells which collectively use the complete set of available frequencies is called a Cluster. The value of N is calculated by the following formula:

$N = I^2 + I*J + J^2$
Where **I,J** =0,1,2,3….
Hence, possible values of N are 1,3,4,7,9,12,13,16,19 and so on.

If a Cluster is replicated or repeated M times within the cellular system, then Capacity, C, will be,

C = MkN = MS

CHANNEL ASSIGNMENT STRATEGIES-

There are two channel assignment strategies in cellular system.

A. Fixed channel assignment:

1. In fixed channel assignment each cell is permanently allocated predetermined group of channels. Any call attempt within cell can only be served by unused channels in that particular cell.
2. If all channels are occupied, the call is blocked and subscriber does not receive service.
3. Borrowing technique where a cell is allowed to borrow channels from a neighboring cell if all channels are already occupied is always used with this type of strategy. Mobile Base station (MSC) monitors the function of base station including borrowing ensuring that borrowing does not interfere with any call in progress in donor cell.

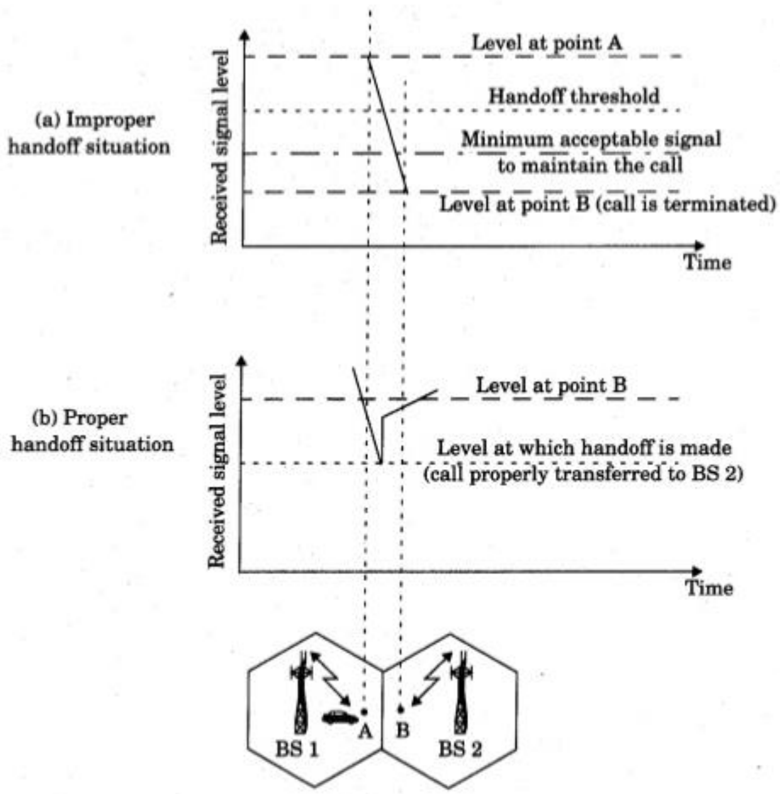
B. Dynamic channel assignment:

1. In dynamic channel assignment strategy, voice channels are not allocated permanently.
2. Entire pool of frequency channels lies with MSC and each time a call request is made, the serving base station requests a channel from the MSC. Switch then allocates a channel to the requested cell following a algorithm.
3. MSC allocates frequency channels on dynamic basis if that frequency channel is not presently in use in the cell or any other cell which falls within the minimum restricted distance of frequency reuse to avoid co-channel interference.
4. It reduces chances of blocking which increases trunking capacity of system as all available channels are accessible to all cells.
5. In this MSC has to collect real time data on channel occupancy, traffic distribution, radio signal strength indication of all channels on continuous basis, thus increasing the computational load on MSC.

HANDOFF-

When a mobile moves into a different cell while a conversation is in progress, the MSC automatically transfers the call to a new channel belonging to the new base station. This handoff operation not only involves identifying a new base station, but also requires that the voice and control signals be allocated to channels associated with the new base station.

Processing handoffs is an important task in any cellular radio system. Many handoff strategies prioritize handoff requests over call initiation requests when allocating unused channels in a cell site. System designers must specify an optimum signallevel at which to initiate a handoff. Once a particular signal level is specified as the minimum usable signal for acceptable voice quality at the base station receiver, a slightly stronger signal level is used as a threshold at which a handoff is made.



This margin, given by $\Delta = P_{r\ handoff} - P_{r\ minimum\ usable}$, cannot be too large or too small. If Δ is too large, unnecessary handoffs which burden the MSC may occur, and if Δ is too small, there may be insufficient time to complete a handoff before a call is lost due to weak signal conditions.

The time over which a call may be maintained within a cell, without handoff, is called the***dwell time.***

In second generation systems, handoff decisions are *mobile assisted*. In ***mobile assisted handoff*** **(MAHO)**, every mobile station measures the received power from surrounding basestations and continually reports the results of these measurements to the serving base station. A handoff is initiated when the power received from the base station of a neighboring cell begins toexceed the power received from the current base station by a certain level or for a certain period of time. The

MAHO method enables the call to be handed over between base stations at a much fasterrate than in first generation analog systems since the handoff measurements are made by each mobile, and the MSC no longer constantly monitors signal strengths.

if a mobile moves from one cellular system to a different cellu-lar system controlled by a different MSC, an *intersystem handoff* becomes necessary.

Prioritizing Handoffs

1. One method for giving priority to handoffs is called the *guard channel concept*, whereby a fraction of the total available channels in a cell is reserved exclusively for handoff requests from ongoing calls which may be handed off into the cell.
2. Queuing of handoff requests is another method to decrease the probability of forced termination of a call due to lack of available channels.

Types of Handoff:

1. **Hard Handoff:**

When there is an actual break in the connectivity while switching from one Base Station to another Base Station. There is no burden on the Base Station and MSC because the switching takes place so quickly that it can hardly be noticed by the users. The connection quality is not that good. Hard Handoff adopted the 'break before make' policy.

2. **Soft Handoff:**

In Soft Handoff, at least one of the links is kept when radio signals are added or removed to the Base Station. Soft Handoff adopted the 'make before break' policy. Soft Handoff is more costly than Hard Handoff.

CO-CHANNEL INTERFERENCE AND SYSTEM CAPACITY OF A CELLULAR RADIO SYSTEMS-

Frequency reuse implies that in a given coverage area there are several cells that use the same setof frequencies. These cells are called *co-channel cells*, and the interference between signals from these cells is called *co-channel interference*.

Thermal noise which can be overcomeby increasing the signal-to-noise ratio (SNR), co-channel interference cannot be combated by simply increasing the carrier power of a transmitter.

This is because an increase in carrier transmit power increases the interference to neighboring co-channel cells. To reduce co-channel interference, co-channel cells must be physically separated by a minimum distance to provide sufficient isolation due to propagation.
When the size of each cell is approximately the same and the base stations transmit the same power, the co-channel interference ratio is independent of the transmitted power and becomes a function of the radius of the cell (R) and the distance between centers of the nearest co-channel cells (D).

$$D = R * (3 * N)^{1/2}$$

Where,
R = Radius of a cell
N = Number of cells in a given cluster

By increasing the ratio of D/R, the spatial separation between co-channel cells relative to the coverage distance of a cell is increased. Thus, interference is reduced from improved isolation of RF energy from the co-channel cell. The parameter Q, called the *co-channel reuse ratio*, is related to the cluster size. For a hexagonal geometry-

$$Q=\frac{D}{R}=\sqrt{3N}$$

A small value of Q provides larger capacity since the cluster size N is small, whereas a large value of Q improves the transmission quality, due to a smaller level of co-channel interference. A trade-off must be made between these two objectives in actual cellular design.

Cluster Size (N)		**Co-channel Reuse Ratio (Q)**
$i=1, j=1$	3	3
$i=1, j=2$	7	4.58
$i=2, j=2$	12	6
$i=1, j=3$	13	6.24

Let i_0 be the number of co-channel interfering cells. Then, the signal-to-interference ratio (S/I or SIR) for a mobile receiver which monitors a forward channel can be expressed as

$$\frac{S}{I}=\frac{S}{\sum_{i=I}^{i_o} I_i}$$

Where S is the desired signal power from the desired base station

I_i is the interference powercaused by the i^{th} interfering co-channel cell base station.

The average received power P_r at a distance d from the transmitting antenna is approximated by

$$P_r=P_0(\frac{d}{d_0})^{-n}$$

$$\text{Or } P_r(dBm)=P_0(dBm)-10nlog(\frac{d}{d_o})$$

Where P_0 is the power received at a close-in reference point in the far field region of the antenna at asmall distance d_0 from the transmitting antenna and n is the path loss exponent.

When the transmit power of each base station is equal and the path loss exponent is thesame throughout the coverage area, S/I for a mobile can be approximated as

$$\frac{S}{I} = \frac{R^{-n}}{\sum_{i=I}^{i_o} (D_i)^{-n}}$$

Considering only the first layer of interfering cells, if all the interfering base stations are equidistant from the desired base station and if this distance is equal to the distance D between cell centers.

$$\frac{S}{I} = \frac{({}^{D}/_{R})^{n}}{i_0} = \frac{(\sqrt{3N})^{n}}{i_0}$$

The above equation relates S/I to the cluster size N, which in turn determines the overall capac ity of the system.

For the U.S. AMPS cellular system which uses FM and 30 kHz channels, subjective tests indicate that sufficient voice quality is provided when S/I is greater than or equal to 18 dB

CONCEPT OF IMPROVING COVERAGE AND CAPACITY IN CELLULAR SYSTEM-

There is a performance criterion of cellular mobile systems like:

a) Voice quality

b) Service Quality like coverage and quality of service.

c) Number of Dropped calls.

d) Special features like call forwarding, call diverting, call barring.

As the demand for wireless service increases, the number of channels assigned to cell becomes insufficient to support required number of users.

At this point, cellular design techniques are needed to provide more channels per unit coverage area.

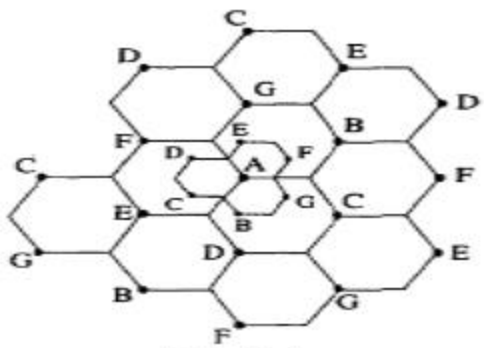
There are 3 techniques for improving cell capacity in cellular system, namely:

- Cell Splitting.
- Sectoring.
- Coverage Zone Approach.

A) **CELL SPLITTING:**

- It is process of subdividing a congested cell into smaller cells, each with its own base station and a corresponding reduction in antenna height and transmitter power.
- Cell splitting increases capacity of cellular system since it increases number of times that channels are reused, it preserves frequency reuse plan.
- It defines new cells which have smaller radius than original cells and by installing these smaller cells called microcells between existing cells, which radius will be half of the original cell.

- Thus capacity increases due to additional number of channels per unit area, but does not disturb the channel allocation scheme required to maintain the minimum co-channel reuse ratio Q between co-channel cells.

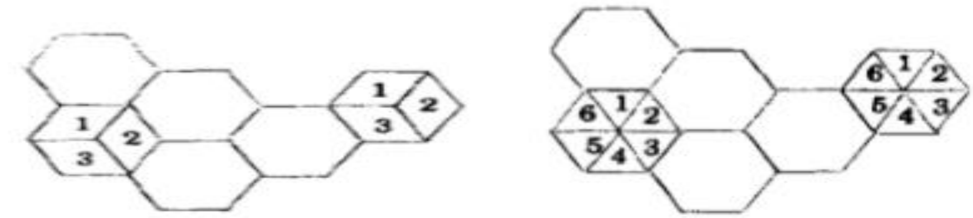


Cell splitting.

B) **SECTORING:**

- This is another method to increase cellular capacity and coverage by keeping cell radius unchanged and decreasing D/R ratio.
- In this approach, capacity improvement is achieved by reducing the number of cells in a cluster and thus increasing the frequency reuse.
- The co-channel interference in a cellular system may be decreased by replacing a single Omni-directional antenna at the base station by several directional antennas, each radiating within a specified sector.
- The factor by which the co-channel interference is reduced depends on the amount of sectoring used.

a) 1200 sectoring b) 600 sectoring



a) 120° sectoring b) 60° sectoring

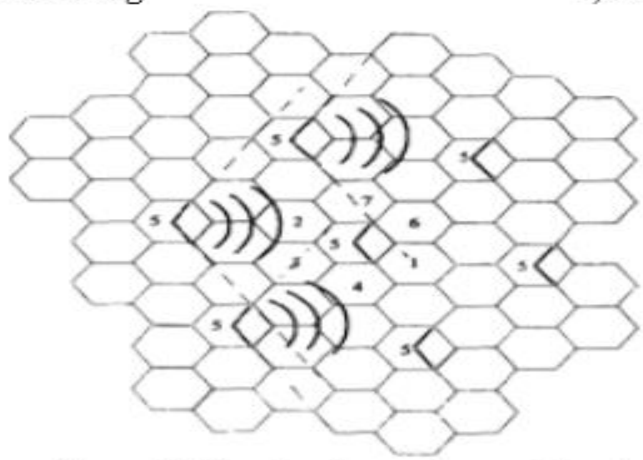


Illustration of how 120° sectoring reduces interference from

Advantages:

- Improvement in Signal capacity.
- Improvement in signal to interference ratio.
- Increases frequency reuse.

Disadvantages:

- Increase in number of handoffs.
- Increase in number of antenna at each base station.

WIRELESS SYSTEMS AND ITS STANDARDS-

The most common set of wireless standard is the IEEE 802.11 Wireless LAN (WLAN) & Mesh. The IEEE updates the 802.11 Wi-Fi standard every few years.

802.1x

This standard enhances the security of local area networks by providing an authentication framework allowing users to authenticate to a central authority, such as LDAP or Active Directory. In conjunction with 802.11 access technologies, it provides an effective mechanism for controlling access to the wireless local area network.

802.11a

An extension to the 802.11 standard developed by the IEEE for wireless network technology. 802.11a applies to wireless local area networks and supports a maximum a maximum connect rate of 54 Mbps throughput in the 5GHz band. This specification is not backwardly compatible with 802.11b/g and requires special wireless adapters.

802.11b

An extension to the 802.11 standard developed by the IEEE for wireless network technology. 802.11b applies to wireless local area networks and supports a maximum connect rate of 11 Mbps with fallback to 5.5, 2, and 1 Mbps in the 2.4GHz ISM band. This standard was ratified in 1999.

802.11g

An extension to the 802.11 standard that allows for a maximum connect rate of 54 Mbps while maintaining compatibility with the 802.11b standard in the 2.4GHz band This specification is compatible and complimentary to the 802.11b standard.

802.11i

An extension to the 802.11 standard to provide improved security over that which is available under 802.11 extensions. This extension provides for improved encryption methods and for the integration of the IEEE 802.1x authentication protocol as well as advanced encryption mechanisms such as AES (Advanced Encryption Standard), for an optional, fully compliant implementation of 802.11i.

802.11n

Uses multiple transmitter and receiver antennas (also known as multiple-input and multiple-output, or MIMO) to allow for increased data throughput and range. This standard was ratified in 2009. Pre-standard hardware is commercially available and not compatible with PittNet Wi-Fi.

GSM SERVICE AND FREATURES-

GSM services are classified as either teleservices or data services. Teleservices include standard mobile telephony and mobile-originated traffic. Data services include computer to computer communication and packet switched traffic. User services may be divided into three major categories.

A. Telephone services: These include emergency calling and facsimile. GSM also supports Videotext and Teletext.

B. Bearer services or data services: These are limited to layer 1,2 and 3 of the open system interconnection (OSI) reference model. Supported services include packet switched protocols and data rates from 300bps to 9.6 kbps. Data may be transmitted using transparent or non-transparent mode.

C. Supplementary ISDN services: These are digital in nature and include call diversion, closed user groups and caller identification, and are not available in analog mobile networks. Supplementary services also include short messaging service (SMS) which allows GSM subscribers and base station to transmit alphanumeric pages of limited length while simultaneously carrying normal voice traffic. SMS provides cell broadcast also can be used for safety and advisory applications such as the broadcast of highway or weather information to all GSM subscribers.

The features of GSM are:

A. Subscriber Identity Module (SIM):

- It is a memory device that stores information such as subscriber's identification number. The networks and countries where the subscriber is entitled to service, privacy keys and other user specific information.
- A subscriber uses SIM with four digit personal ID number to activate service from any GSM phone.
- SIM is available as smart card or plug in module, which is less convenient than the SIM card but is removable and portable.
- Without SIM installed all GSM mobiles are identical and non-operational. SIM gives GSM subscriber units their identity.
- Subscriber can plug their SIM into any suitable terminal such as a hotel phone, public phone , or any portable or mobile and are then able to have all incoming GSM calls routed to that terminal and have all outgoing calls billed to their home phone no matter where they are in the world.

B. On the air privacy:

- Unlike analog FM cellular phone system which can be readily monitored, it is virtually impossible to eavesdrop on a GSM radio transmission.
- The privacy is made possible by encrypting the digital bit stream sent by a GSM transmitter, according to a specific secret cryptographic key that is known only to the cellular carrier. This key changes with time for each user.
- Every carrier and GSM equipment manufacturer has to sign an international agreement MoU which allows the sharing of cryptographic algorithms and other proprietary information between countries and carriers.

ARCHITECTURE OF GSM SYSTEM AND GSM MOBILE STATION-

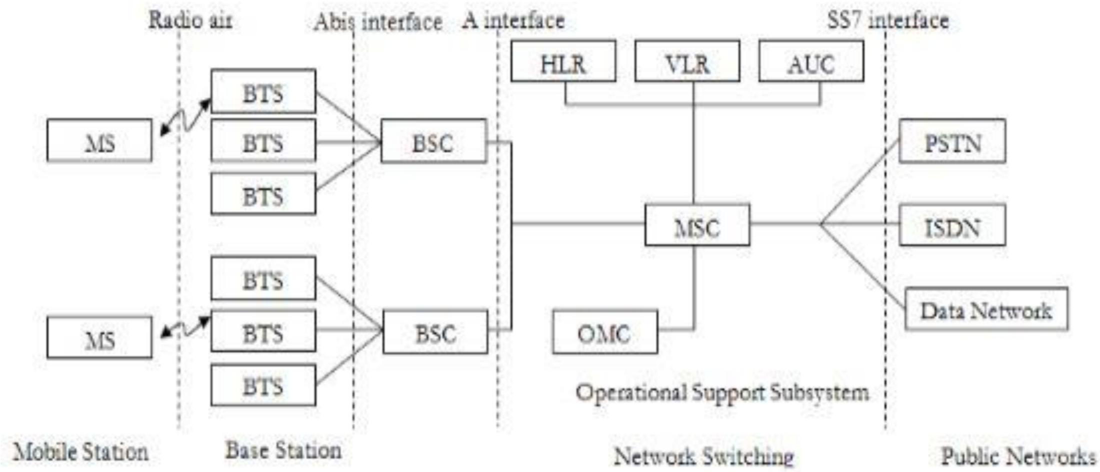


Fig: GSM Architecture

The GSM architecture consists of three major interconnected subsystems that interact with themselves and with users through certain network interface. The subsystems are Base Station Subsystem (BSS), Network Switching Subsystem (NSS) and Operational Support Subsystem (OSS). Mobile Station (MS) is also a subsystem but it is considered as a part of BSS.

1. Mobile Station (MS): Mobile Station is made up of two entities.

A. Mobile equipment (ME):

- It is a portable, vehicle mounted, hand held device.
- It is uniquely identified by an IMEI number.
- It is used for voice and data transmission. It also monitors power and signal quality of surrounding cells foe optimum handover. 160 characters long SMS can also be sent using Mobile Equipment.

B. Subscriber Identity module (SIM):

- It is a smart card that contains the International Mobile Subscriber Identity (IMSI) number.
- It allows users to send and receive calls and receive other subscriber services. - It is protected by password or PIN.
- It contains encoded network identification details. it has key information to activate the phone.
- It can be moved from one mobile to another.

2. Base Station Subsystem (BSS): It is also known as radio subsystem, provides and manages radio transmission paths between the mobile station and the Mobile Switching Centre (MSC). BSS also manages interface between the mobile station and all other subsystems of GSM. It consists of two parts.

A. Base Transceiver Station (BTS):

- It encodes, encrypts, multiplexes, modulates and feeds the RF signal to the antenna.
- It consists of transceiver units.
- It communicates with mobile stations via radio air interface and also communicates with BSC via Abis interface.

B. Base Station Controller (BSC):

- It manages radio resources for BTS. It assigns frequency and time slots for all mobile stations in its area.
- It handles call set up, transcoding and adaptation functionality handover for each MS radio power control.
- It communicates with MSC via A interface and also with BTS.

3. Network Switching Subsystem (NSS): it manages the switching functions of the system and allows MSCs to communicate with other networks such as PSTN and ISDN. It consist of

A. Mobile switching Centre:

- It is a heart of the network. It manages communication between GSM and other networks.
- It manages call set up function, routing and basic switching.
- It performs mobility management including registration, location updating and inter BSS and inter MSC call handoff.
- It provides billing information.
- MSC does gateway function while its customers roam to other network by using HLR/VLR.

B. Home Location Registers (HLR): - It is a permanent database about mobile subscriber in a large service area. - Its database contains IMSI, IMSISDN, prepaid/post-paid, roaming restrictions, supplementary services.

C. Visitor Location Registers (VLR): - It is a temporary database which updates whenever new MS enters its area by HLR database. - It controls mobiles roaming in its area. It reduces number of queries to HLR. - Its database contains IMSI, TMSI, IMSISDN, MSRN, location, area authentication key.

D. Authentication Centre: - It provides protection against intruders in air interface. - It maintains authentication keys and algorithms and provides security triplets (RAND, SRES, Ki).

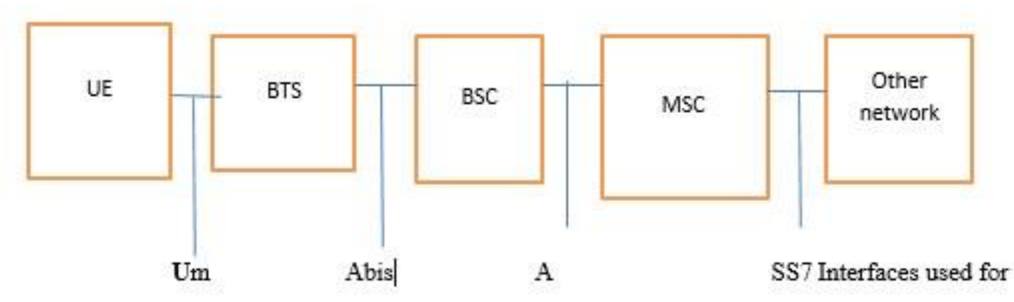
E. Equipment Identity Registry (EIR):

- It is a database that is used to track handset using the IMEI number.
- It is made up of three sub classes- the white list, the black list and the gray list.

4. Operational Support Subsystem (OSS): It supports the operation and maintenance of GSM and allows system engineers to monitor, diagnose and troubleshoot all aspects of GSM system. It supports one or more Operation Maintenance Centres (OMC) which are used to monitor the performance of each MS, Bs, BSC and MSC within a GSM system. It has three main functions:

- To maintain all telecommunication hardware and network operations with a particular market.
- To manage all charging and billing procedures
- To manage all mobile equipment in the system.

Interfaces used for GSM network :
1)UM Interface –Used to communicate between BTS with MS
2)Abis Interface— Used to communicate BSC TO BTS
3)A Interface-- Used to communicate BSC and MSC
4) Singling protocol (SS 7)- Used to communicate MSC with other network .



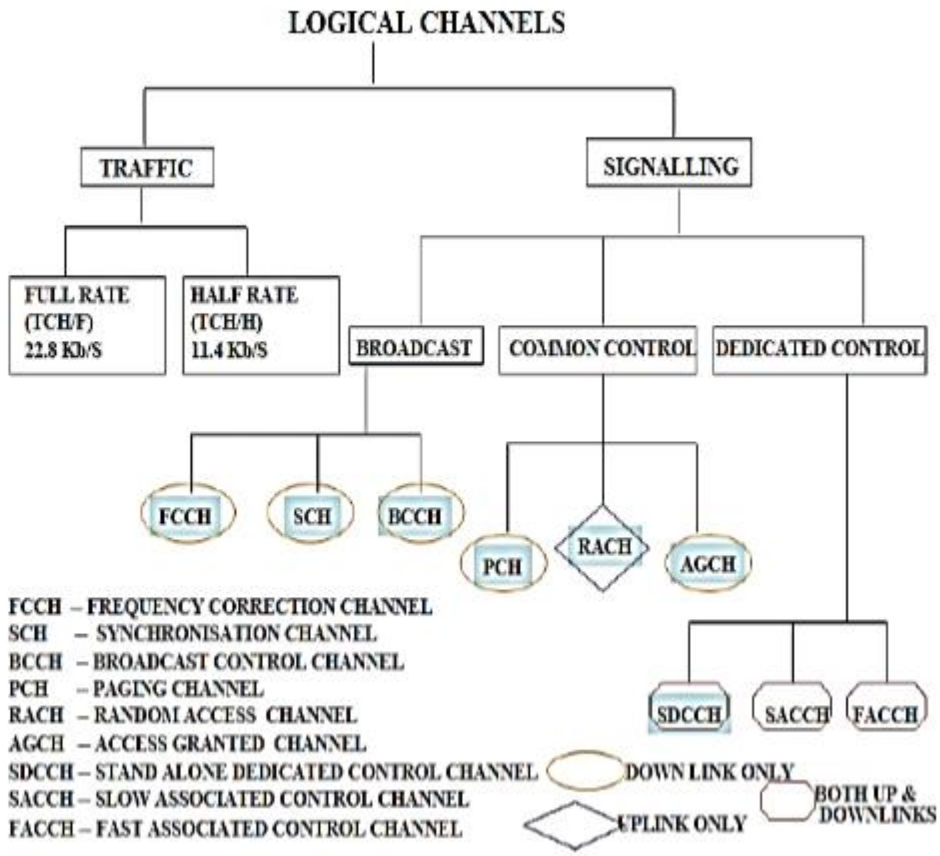
CHANNEL TYPES OF GSM SYSTEM-

GSM Channel are divided into two types:

Traffic channels (TCHs):

- The traffic channels are intended to carry encoded speech or user data.
- Traffic channels are intended to carry encoded speech and user data.
- Full rate traffic channels at a net bit rate of 22.8 Kb/s (TCH/F)
- Half rate traffic channels at a net bit rate of 11.4 Kb/s (TCH/H)
- Speech channels are defined for both full rate and half rate traffic channels.

- Data channels support a variety of data rates (2.4, 4.8 and 9.6 Kb/s) on both half and full rate traffic channels. The 9.6 Kb/s data rate is only for full rate application.



- The traffic channels(TCH) support two types of information rates Full rate (TCH/F) and Half rate (TCH/H)
- When transmitted as full rate, the user data is occupied within TS per frame. When transmitted as half rate, the user data is occupied into the same time slot but sent in alternate frames.
- The 26th frame contains idle bits if full rate TCHs are used and contains SACCH data if half rate TCHS are used

Full Rate TCH for data and speech channels:-

A. **Full - rate Speech Channel) TCH/Fs):** At 16 kbps the full rate speech channel is digitized. The full rate speech channel caries 55.8kbps after adding the GSM channel coding to the digitized speech.

B. **Full-rate Data Channel for 9600 bps (TCH/F9.6):** The full rate traffic data channel contains raw data that is transmitted at 9.6 kbps. After the application of additional forward error correction coding with the GSM standards, 9600 kbps is transferred at 22.8 kbps.

C. **Full-rate Data Channel for 4500 bps (TCH/F4.8):** The full rate traffic date channel contains data that is transmitted at 4.8 Kbps. After the application of additional forward error correction coding with GSM standards, the 4.8 kbps is transferred at 22.8 kbps.

D. **Full Rate Data Channel for 2400 bps (TCH/F2.4):** The full rate traffic data channel contains raw data that is transmitted at 2.4 kbps. After the application of additional forward error correction coding with GSM standards, the 2.4 kbps data is transferred at 22.8 kbps.

Half Rate TCH for data and speech channels:

A. **Half Rate Speech Channels (TCH/HS):** The half tate speech channel can carry digitized speech that is sampled at a rate half that of full rate channel. GSM anticipates the availability of speech coders. It can digitize speech at about 6.5 kbps. After adding GSM channel coding to the digitized speech, the half rate Speech channel will carry 11.4 kbps.

B. **Half Rate Data Channel for 4800 bps (TCH/H4.8):** The half rate traffic data channel carries raw data that is sent at 4800 bps. After the application of forward error correction using GSM standards, 4800 bps data is sent at 11.4 kbps.

C **Half Rate Data Channel for 2400 kbps (TCH/H 2.4):** The half rate traffic data channel carries raw user data that is sent at 2400 bps. After application of additional forward error correction using GSM standards, 2400 bps data is sent to 11.4 bps.

Control Channel (CCH):

. Control channels carry signaling information between an MS and a BTS.

a) **Broadcast control channel:**

. Broadcast control channels are transmitted in downlink direction only i.e. only transmitted by BTS.

. The broadcast channels are used to broadcast synchronization and general network information to all the MSs within a cell.

. **It has three types:**

a. **FREQUENCY CORRECTION CHANNEL (FCCH):**

. Used for the frequency correction / synchronization of a mobile station.

. The repeated (every 10 sec) transmission of Frequency Bursts is called FCCH.

b. **SYNCHRONISATION CHANNEL (SCH):**

. Allows the mobile station to synchronize time wise with the BTS.

. Repeated broadcast (every 10 frames) of Synchronization Bursts is called (SCH).

c. **BROADCAST CONTROL CHANNEL (BCH):**

. The BROADCAST CONTROL CHANNEL (BCCH) is used to broadcast control information to every MS within a cell.

. This information includes details of the control channel configuration used at the BTS, a list of the BCCH carrier frequencies used at the neighboring BTSs and a number of parameters that are used by the MS when accessing the BTS.

Common Control Channel:

The common control channels are used by an MS during the paging and access procedures. Common control channels are of three types.

(PCH) PAGING CHANNEL:

. Within certain time intervals the MS will listen to the Paging channel, PCH, to see if the network wants to get in contact with the MS.

. The reason could be an incoming call or an incoming Short Message.

2. **(RACH) RANDOM ACCESS CHANNEL:**

- If listening to the PCH, the MS will realize it is being paged.
- The MS answers, requesting a signaling channel, on the Random Access channel, RACH.
- RACH can also be used if the MS wants to get in contact with the network, e/g. when setting up a mobile originated call.

3. **(AGCH) ACCESS GRANTED CHANNEL:**

- The access grant channel (AGCH) is carried data which instructs the mobile to operate in a particular physical channel (Time slot or ARFCN).
- It uses normal burst.

C) Dedicated Control Channels (DCCHs):

- Signaling information is carried between an MS and a BTS using associated and dedicated control channels during or not during a call, They are of three types:

A. (SDCCH) STAND ALONE DEDICATED CONTROL CHANNEL:

- Non-urgent information, e.g. transmitter power control, is transmitted using the slow associated control channel (SACCH).
- On the uplink MS sends averaged measurements on own base station (signal strength and quality) and neighboring base stations (signal strength).
- On the downlink the MS receives system information, which transmitting power and what timing advance to use. It is transmitted at 13thFrame of TCH. As seen, SACCH is transmitted on both up-and downlink, point-to-point.
- It uses normal burst.

B. (SAACH) SLOW ASSOCIATED CONTROL CHANNEL:

- In some situations, signaling information must flow between a network and an MS when a call is not in progress, e.g. during a location update.
- This could be accommodated by allocating either a full-rate or half-rate TCH and by using either the SACCH or FACCH to carry the information.

C. (FACCH) FAST ASSOCIATED CONTROL CHANNEL:

- More urgent information, e.g. a handover command, is sent using time slots that are 'stolen' from the traffic channel.
- If, suddenly, during the conversation a handover must be performed the Fast Associated Control channel, FACCH, is used.
- FACCH works in stealing mode, meaning that one 2. ms segment of speech is exchanged for signaling information necessary for the handover.

WORKING OF FORWARD AND REVERSE CDMA CHANNEL-

CDMA channels can be broadly categorized as Forward channel and Reverse channel.

Forward Channel

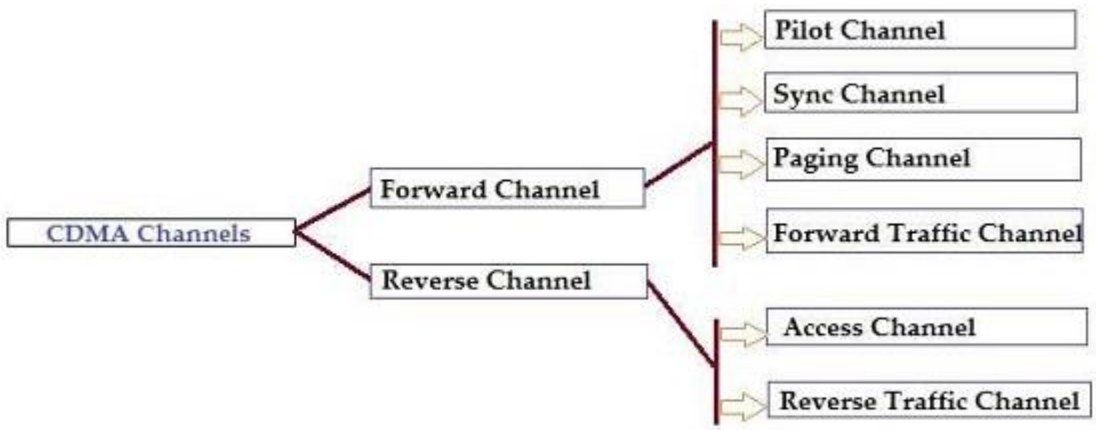
The Forward channel is the direction of the communication or mobile-to-cell downlink path. It includes the following channels −

- **Pilot Channel** − Pilot channel is a reference channel. It uses the mobile station to acquire the time and as a phase reference for coherent demodulation. It is continuously transmitted by each base station on each active CDMA frequency. And, each mobile station tracks this signal continuously.
- **Sync Channel** − Synchronization channel carries a single, repeating message, which gives the information about the time and system configuration to the mobile station. Likewise, the mobile station can have the exact system time by the means of synchronizing to the short code.
- **Paging Channel** − Paging Channel's main objective is to send out pages, that is, notifications of incoming calls, to the mobile stations. The base station uses these pages to transmit system overhead information and mobile station specific messages.
- **Forward Traffic Channel** – Forward Traffic Channels are code channels. It is used to assign calls, usually voice and signaling traffic to the individual users.

Reverse Channel

The Reverse channel is the mobile-to-cell direction of communication or the uplink path. It consists of the following channels −

- **Access Channel –** Access channel is used by mobile stations to establish a communication with the base station or to answer Paging Channel messages. The access channel is used for short signaling message exchanges such as call-ups, responses to pages and registrations.
- **Reverse Traffic Channel –** Reverse traffic channel is used by the individual users in their actual calls to transmit traffic from a single mobile station to one or more base stations.



FREQUENCY AND CHANNEL SPECIFICATION-

This channel may be voice data or overhead control data. A CDMA channel is a pair of 1.25MHz frequency bands separated by guard-band. The guard-band may be 45 or 80MHz wide.

The pair of 1.25MHz band is located either in the 82-893 MHz range in case of cellular CDMA or in the 1850-1989 MHz range in case of PCS CDMA.

The guard band used in cellular is 45MHz for PCS for PCS it is 80 MHZ. In both the cellular and PCS frequency ranges, CDMA channels have been identified by assigning CDMA channel numbers to the associated center frequencies for the two 1.25 MHz bands used by the channel.

The channel numbers lie in the range 0-1023. Out of the two 1.25MHz bands associated with a CDMA channel, the higher frequency one is called the forward CDMA and the lower frequency one the reverse CDMA channel. The terms forward and reverse link are also used to refer to them.

The forward link is used for any communication directed by the base station towards mobiles located within its coverage area.

The Reverse link is used for any communication directed by the mobiles towards the base station with which they are connected .In general, each cell is assigned a specific CDMA channel which must be used by its base station and the mobiles within the cell. If the base station uses a sectorized antenna, each sector may be assigned a different CDMA channel. Thus sectorized cells may have 3 or 6 channels, instead of just one.

ARCHITECTURE AND FEATURES OF GPRS-

General packet Radio Service (GPRS) is a enhancement of GPS, which is packet oriented mobile data service on the 2G and 3G cellular communication system. It provide connection to the external packet data network through the GSM infrastructure with short access time to the network for independent short packets.

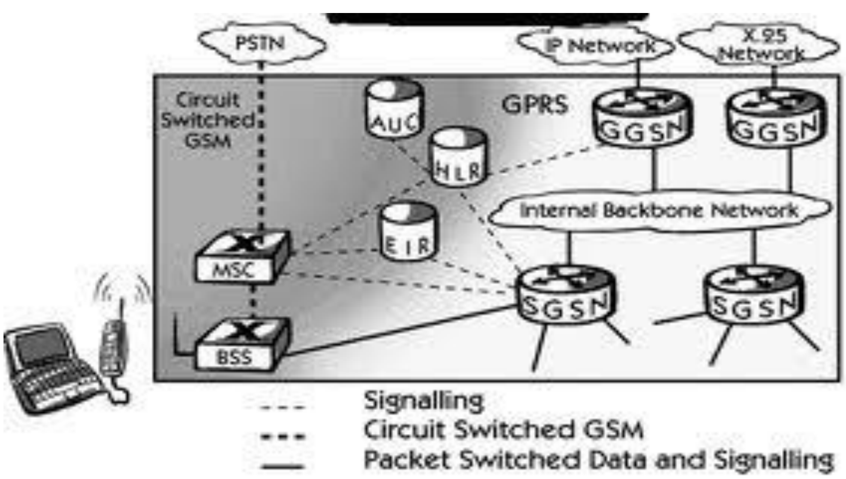
It uses exactly uses the same physical radio channel as GSM and only new logical GPRS Radio Channel are defined. GPRS was originally standardized by European Telecommunications Standards Institute (ETSI) in response to the earlier CDPD & i-mode packet-switched cellular technologies. Here in this post, GPRS Architecture in Mobile Communication is explained in detail.

GPRS Network Architecture:

GPRS is usually attempts to reuse the existing GSM network elements as much as possible. There are new entities called GPRS that supports nodes (GSN) which are responsible for delivery and routing of data packets between mobile stations and external packets networks. There are two types of GSNs,

- Serving GPRS Support Node (SGNS)
- Gateway GPRS Support Node (GGNS)

These two modes are comparable to MD-IS in CPDP. There is also a new database called GPRS register which is located with HLR. It stores routing information and maps the IMSI to a PDN address. Thus, GPRS Reference Architecture is shown as-



GPRS Network Architecture

Subsystems of GPRS Architecture

Mobile Station:

GPRS Services required New Mobile Station as the existing GSM phones are not capable of handling the enhanced air interface or the packet data. A wide variety of Mobile stations exist which includes a high-speed version of current phones to support high-speed data access like PC cards for laptop computers. These mobile stations are in backward compatibility mode in order to make voice calls which are used GSM.

Base Station Subsystem:

Each BSC requires the installation of Packet Control Units in addition to software upgrade. They provide physical and logical data interface to BSS to estimate packet data traffic. BTS too require a software upgrade but typically does not involve hardware enhancements.

When the traffic is originated at the subscriber mobile then it is transported over the air interface to BTS and then from BTS to BSC, the same way in standard GSM call. But at output of BSC the traffic is separated, the voice is sent to the mobile switching centre per standard GSM and the data is sent to the new device called the SGSN via the PCU.

GRPS Support Nodes:

- **SSGN:** The Serving GPRS Support Node is responsible for authentication of GPRS mobiles, registration of mobiles in the network, mobility management, and collecting information for charging for the use of the air interface.
- **GGSN:** The Gateway GPRS Support Node acts as an interface and a router to external networks. The GGSN contains routing information for GPRS mobiles, which is used to tunnel packets through the IP based internal backbone to the correct Serving GPRS Support Node.

Internal Back Network:

The internal backbone is an IP based network which is used to carry the new packets between different GSN. The process of Tunneling is used in-between SGSNs and GGSNs, this is done to safe exchange of domain information outside the GPRS Network without informing internal backbone.

Mobility Support:

In a manner similar to GSM and CDPD, there are mechanism in GPRS to support mobility. There are two types of Mobility Support in GPRS Network-

- Attachment Procedure
- Location and Handoff Management

Short Messaging Services in GSM:

For the proliferation of GSM enable the introduction of SMS, which is similar to peer-to-peer instant messaging on the Internet. Users of SMS can exchange alphanumeric message of up to 160 char. within seconds of submission of the message.

MOBILE TCP PROTOCOL –

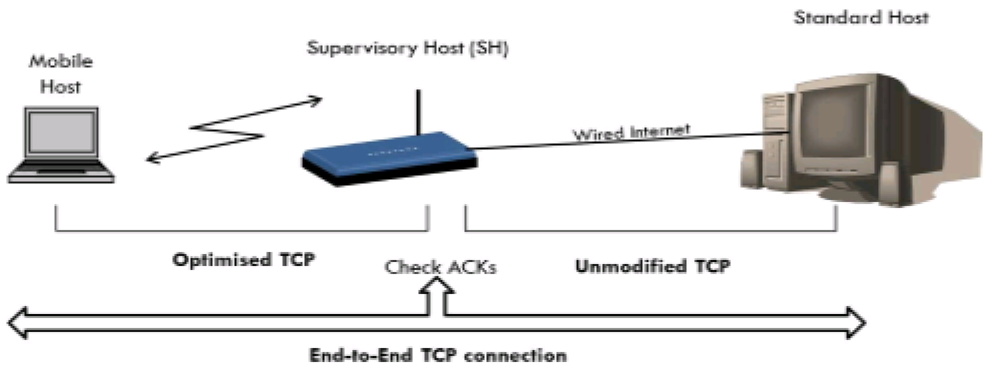
- The M-TCP splits up the connection into two parts:
  - An unmodified TCP is used on the Standard host-Supervisory Host section
  - An optimized TCP is used on the Supervisory Host- Mobile Host section.
- The Supervisory Host (SH) adorns the same role as the proxy (Foreign Agent) in I-TCP.
- The SH is responsible for exchanging data to both the Standard host and the Mobile host.
- Here in this approach, we assume that the error bit rate is less as compared to other wireless links.
- So if any packet is lost, the retransmission has to occur from the original sender and not by the SH. (This also maintains the end-to-end TCP semantic)
- The SH monitors the ACKs (ACK means acknowledgement) being sent by the MH. If for a long period ACKs have not been received, then the SH assumes that the MH has been disconnected (maybe due to failure or moved out of range, etc...).
- If so the SH chokes the sender by setting its window size to 0.
- Because of this the sender goes into persistent mode i.e. the sender's state will not change no matter how long the receiver is disconnected.
- This means that the sender will not try to retransmit the data.
- Now when the SH detects a connectivity established again with the MH (the old SH or new SH if handover), the window of the sender is restored to original value.

Advantages:

- Maintains the TCP end-to-end semantics. (No failed packet retransmission is done by the SH .All job handled by original sender)
- Does not require the change in the sender's TCP.
- If MH disconnected, it doesn't waste time in useless transmissions and shrinks the window size to 0.
- No need to send old buffer data to new SH in case of handover (as in I-TCP).

Disadvantages:

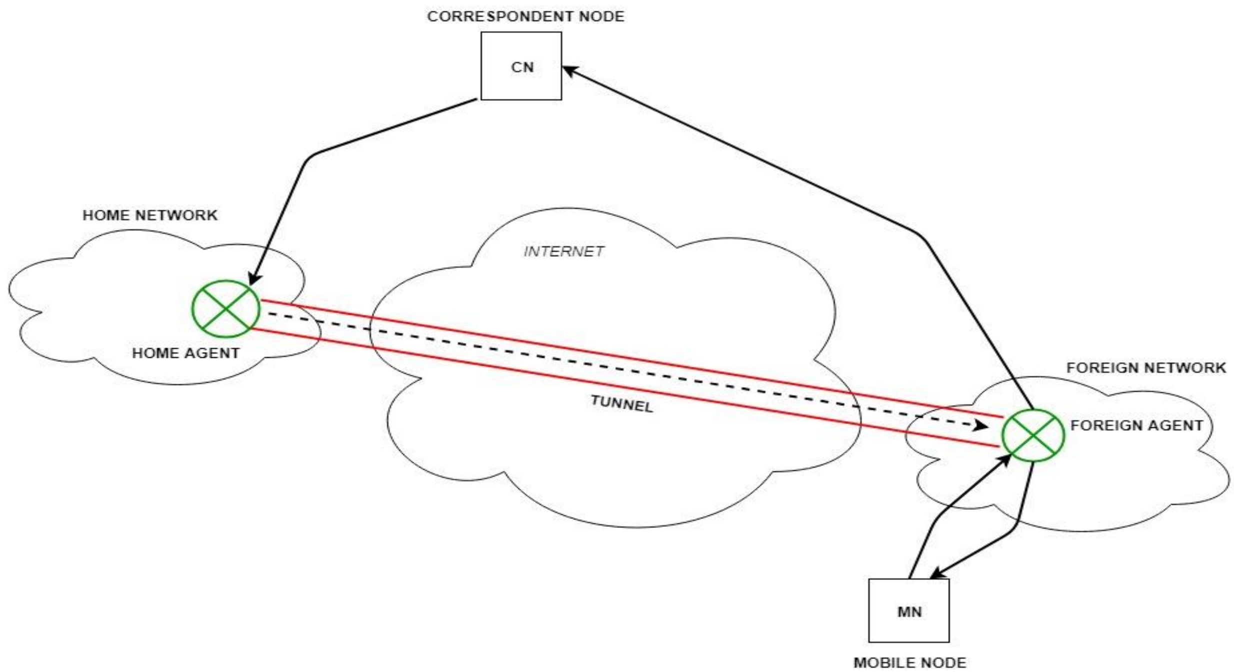
- M-TCP assumes low bit error which is not always true. So, any packet loss due to bit-errors occurring, then its propagated to the sender.
- Modifications are required for the MH protocol software.



MOBILE IP PROTOCOL-

Mobile IP is a communication protocol (created by extending Internet Protocol, IP) that allows the users to move from one network to another with the same IP address. It ensures that the communication will continue without user's sessions or connections being dropped.
Terminologies:

- Mobile Node (MN):
  It is the hand-held communication device that the user caries e.g. Cell phone.
- Home Network:
  It is a network to which the mobile node originally belongs to as per its assigned IP address (home address).
- **Home Agent (HA):**
  It is a router in home network to which the mobile node was originally connected
- **Home Address:**
  It is the permanent IP address assigned to the mobile node (within its home network).
- Foreign Network:
  It is the current network to which the mobile node is visiting (away from its home network).
- **Foreign Agent (FA):**
  It is a router in foreign network to which mobile node is currently connected. The packets from the home agent are sent to the foreign agent which delivers it to the mobile node.
- **Correspondent Node (CN):**
  It is a device on the internet communicating to the mobile node.
- **Care of Address (COA):**
  It is the temporary address used by a mobile node while it is moving away from its home network.



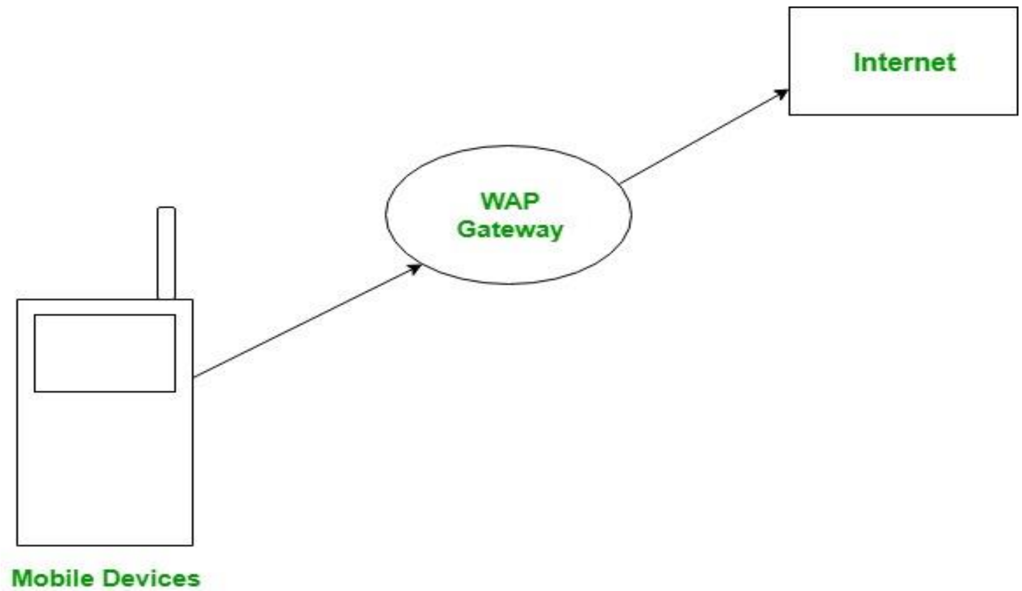
Working:
Correspondent node sends the data to the mobile node. Data packets contains correspondent node's address (Source) and home address (Destination). Packets reaches to the home agent. But now mobile node is not in the home network, it has moved into the foreign network. Foreign agent sends the care-of-address to the home agent to which all the packets should be sent. Now, a tunnel will be established between the home agent and the foreign agent by the process of tunneling.
Tunneling establishes a virtual pipe for the packets available between a tunnel entry and an endpoint. It is the process of sending a packet via a tunnel and it is achieved by a mechanism called encapsulation.

Now, home agent encapsulates the data packets into new packets in which the source address is the home address and destination is the care-of-address and sends it through the tunnel to the foreign agent. Foreign agent, on other side of the tunnel receives the data packets, de-capsulate them and sends them to the mobile node. Mobile node in response to the data packets received, sends a reply in response to foreign agent. Foreign agent directly sends the reply to the correspondent node.

WIRELESS APPLICATION PROTOCOL-

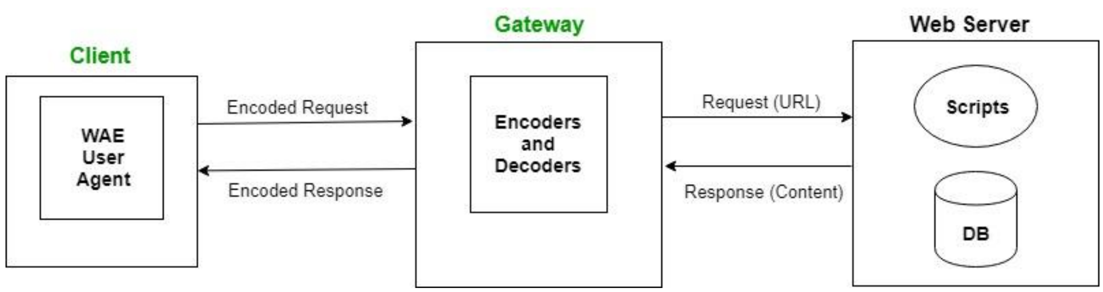
WAP stands for Wireless Application Protocol. It is a protocol designed for micro-browsers and it enables the access of internet in the mobile devices. It uses the mark-up language WML (Wireless Markup Language and not HTML), WML is defined as XML 1.0 application. It enables creating web applications for mobile devices. In 1998, *WAP Forum* was founded by Ericson, Motorola, Nokia and Unwired Planet whose aim was to standardize the various wireless technologies via protocols.

WAP protocol was resulted by the joint efforts of the various members of WAP Forum. In 2002, WAP forum was merged with various other forums of the industry resulting in the formation of Open Mobile Alliance (OMA).



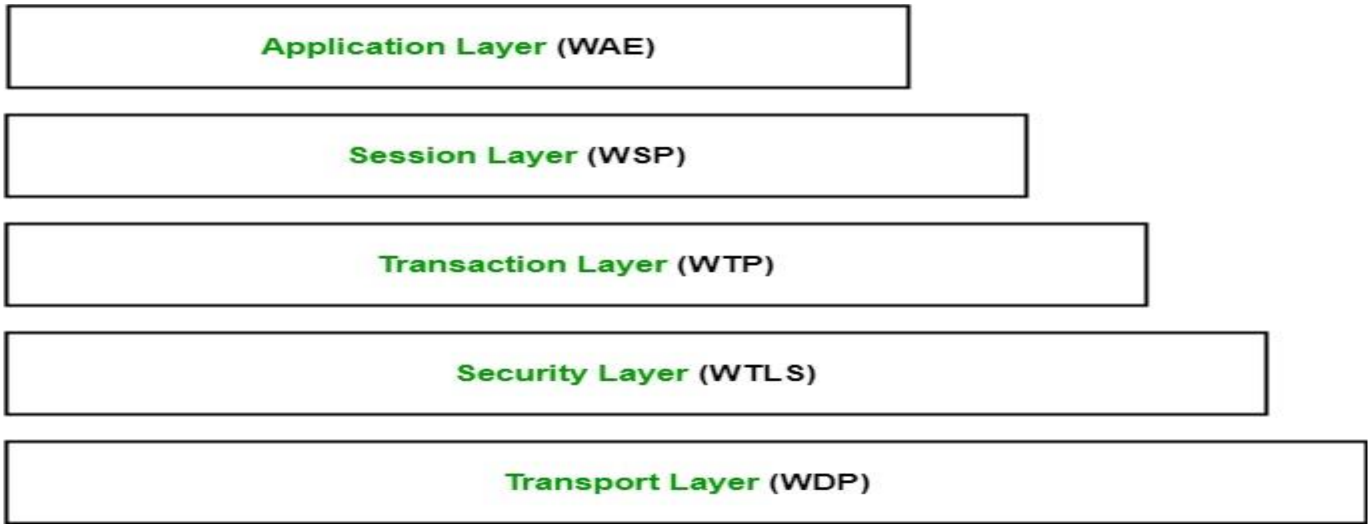
WAP Model:

The user opens the mini-browser in a mobile device. He selects a website that he wants to view. The mobile device sends the URL encoded request via network to a WAP gateway using WAP protocol.



The WAP gateway translates this WAP request into a conventional HTTP URL request and sends it over the internet. The request reaches to a specified Web server and it processes the request just as it would have processed any other request and sends the response back to the mobile device through WAP gateway in WML file which can be seen in the micro-browser.

WAP Protocol stack:



1. Application Layer:
   This layer contains the *Wireless Application Environment (WAE)*. It contains mobile device specifications and content development programming languages like WML.
2. Session Layer:
   This layer contains *Wireless Session Protocol (WSP)*. It provides fast connection suspension and reconnection.
3. Transaction Layer:
   This layer contains *Wireless Transaction Protocol (WTP)*. It runs on top of UDP (User Datagram Protocol) and is a part of TCP/IP and offers transaction support.
4. Security Layer:
   This layer contains *Wireless Transaction Layer Security (WTLS)*. It offers data integrity, privacy and authentication.
5. Transport Layer:
   This layer contains *Wireless Datagram Protocol*. It presents consistent data format to higher layers of WAP protocol stack.

FEATURES OF SMS-

A short message is a brief text message sent to or from a mobile phone subscriber through the Short Message Service (SMS). The standard short message consists of up to 160 alphanumeric characters, although messages at least 50% longer can be sent using data compression. Developed as part of the Global System for Mobile communications (GSM) Phase 1 standard, a short message is exchanged between two mobile devices or between a non-mobile device and a mobile device (for example, a short message can be sent from a PC attached to the Internet to a mobile subscriber). Short messages are stored in and forwarded from a Short Message Service Center (SMSC) so that - unlike the user of a pager - the recipient can get messages that arrive when their mobile device is not turned on.

SMS compression increases the amount of text that can be sent, and SMS concatenation enables short messages to be strung together into a longer one. Users of devices that are not SMS-enabled can send short messages using an alternate version known as *Internet SMS*.

Instant messaging (IM) messages are also sometimes referred to as short messages. To make the most of a short message, people frequently use a shorthand typing mixture of letters and numerals known as Alphanumeric.

MMS

MMS stands for Multimedia Messaging Service. It is the standard way to send messages from one device to another through a network.
As the name Multimedia, we can suggest from here that it is not only for sending text messages, we can also send multimedia like images, audio clips and video clips, and many more things. It is the extension used for SMS(Short Message Service) where we send and receive text messages only with the limitation of only 160 characters in one SMS.
Most of the smartphones support MMS messaging nowadays. Basically it is the advanced version of the text messaging with the additional feature of multimedia.

Modes of sending MMS

There are basically six modes which are as follows:

- Sending messages to an MMS mobile phone via an MMS mobile phone.
  It can be sent in the same way as we send SMS messages, except that MMS messages include multimedia contents.
- Sending messages to a non-MMS mobile phone via an MMS mobile phone.
  Since the non-MMS mobile phones can't receive a multimedia message, the MMS system automatically forwards the messages to the receiver's corresponding email box and then sends a notification to his mobile phone.
- Sending messages to email boxes via an MMS mobile phone Multimedia messages can be sent via an MMS mobile phone to an email box, and the receiver logs on the email box to read the messages. However, most email boxes don't support multimedia messages yet.
- Sending messages to an MMS mobile phone via an email box.
  A user logs on to his email box, selects multimedia messages to send, inputs a receiver's MMS mobile phone number, and send the messages as an attachment.
- Downloading multimedia messages from the internet to an MMS mobile phone.
  A user can customize and order multimedia messages on websites that provide MMSs and then send MMS to an MMS mobile phone.
- Sending messages from an MMS mobile phone to personal e-albums.
  A user can send MMS messages to his personal e-album via an MMS mobile phone. User writes MMS messages in mobile phones, inputs the album website number, and then sends the messages.

FEATURES-

MMS Video Messaging

MMS containing short video clips can be used for a variety of purposes, including sending

subscriber’s movie trailers, advertising businesses, demonstrating new products, or conveying political message.
Symphony provides its users with the opportunity to embed video messages within their marketing campaigns with ease.

MMS Photo Messaging

Photo messaging takes mobile marketing to a new level. Symphony users can quickly attach photos to their outgoing MMS messages for a rich consumer engagement experience. The possibilities are limitless. For example: an online event organizer can send a bar code or QR code directly to event attendees’ phones, as an optional ticket distribution channel.
Now instead of having the attendee print the ticket, they can just show up, show their phone and gain easy access to the event. In short, adding an image to any campaign will boost results. Why tell people about your new dish when you can show it to them? A consumer send address request? Send them back the address coupled with a picture of your business for easy identification.

MMS Mobile Coupon

Businesses can send colorful MMS coupons to customers using MMS picture messaging. Customers can be texted a keyword to receive the coupon or the business can mass send MMS mobile coupons to opt-in customers’ mobile phone. You can offer discounts, buy one get one offers, etc. MMS Coupons or Promotions are messages which are sent to your targeted audiences to offer them the latest products, services and upcoming offers.
With MMS Marketing, mobile marketers have the ability to send rich, quality content to their subscribers. MMS allows for video clips, images or sound file containing enticing mobile content of mobile coupons, product information and more.

1G - FIRST GENERATION

This was the first generation of cell phone technology. The very first generation of commercial cellular network was introduced in the late 70's with fully implemented standards being established throughout the 80's. It was introduced in 1987 by Telecom (known today as Telstra), Australia received its first cellular mobile phone network utilising a 1G analog system. 1G is an analog technology and the phones generally had poor battery life and voice quality was large without much security, and would sometimes experience dropped calls. These are the analog telecommunications standards that were introduced in the 1980s and continued until being replaced by 2G digital telecommunications. The maximum speed of 1G is 2.4 Kbps.

2G - SECOND GENERATION

Cell phones received their first major upgrade when they went from 1G to 2G. The main difference between the two mobile telephone systems (1G and 2G), is that the radio signals used by 1G network are analog, while 2G networks are digital .

Main motive of this generation was to provide secure and reliable communication channel. It implemented the concept of CDMA and GSM. 2G capabilities are achieved by allowing multiple users on a single channel via multiplexing. During 2G Cellular phones are used for data also along with voice.

The advance in technology from 1G to 2G introduced many of the fundamental services that we still use today, such as SMS, internal roaming, conference calls, call hold and billing based on services e.g. charges based on long distance calls and real time billing. The max speed of 2G with General Packet Radio Service (GPRS) is 50 Kbps or 1 Mbps with Enhanced Data Rates for GSM Evolution (EDGE).

Before making the major leap from 2G to 3G wireless networks, the lesser-known 2.5G and 2.75G was an interim standard that bridged the gap.

3G - THIRD GENERATION

Web browsing, email, video downloading, picture sharing and other Smartphone technology were introduced in the third generation. Introduced commercially in 2001, the goals set out for third generation mobile communication were to facilitate greater voice and data capacity, support a wider range of applications, and increase data transmission at a lower cost .

The 3G standard utilizes a new technology called UMTS as its core network architecture - Universal Mobile Telecommunications System. This network combines aspects of the 2G network with some new technology and protocols to deliver a significantly faster data rate. One of requirements set by IMT-2000 was that speed should be at least 200Kbps to call it as 3G service.

3G has Multimedia services support along with streaming are more popular. In 3G, Universal access and portability across different device types are made possible (Telephones, PDA's, etc.). 3G increased the efficiency of frequency spectrum by improving how audio is compressed during a call, so more simultaneous calls can happen in the same frequency range. Like 2G, 3G evolved into 3.5G and 3.75G as more features were introduced in order to bring about 4G.

4G - FOURTH GENERATION

4G is a very different technology as compared to 3G and was made possible practically only because of the advancements in the technology in the last 10 years. Its purpose is to provide high speed, high quality and high capacity to users while improving security and lower the cost of voice and data services, multimedia and internet over IP

Potential and current applications include amended mobile web access, IP telephony, gaming services, high-definition mobile TV, video conferencing, 3D television, and cloud computing.

The key technologies that have made this possible are MIMO (Multiple Input Multiple Output) and OFDM (Orthogonal Frequency Division Multiplexing). The two important 4G standards are WiMAX (has now fizzled out) and LTE (has seen widespread deployment). LTE (Long Term Evolution) is a series of upgrades to existing UMTS technology and will be rolled out on Telstra's existing 1800MHz frequency band.

The max speed of a 4G network when the device is moving is 100 Mbps or 1 Gbps for low mobility communication like when stationary or walking, latency reduced from around 300ms to less than 100ms, and significantly lower congestion.

When 4G first became available, it was simply a little faster than 3G. 4G is not the same as 4G LTE which is very close to meeting the criteria of the standards. To download a new game or stream a TV show in HD, you can do it without buffering.

We have the fractional parts: 4.5G and 4.9G marking the transition of LTE (in the stage called LTE-Advanced Pro) getting us more MIMO, more D2D on the way to IMT-2020 and the requirements of 5G.

5G - FIFTH GENERATION

5G is a generation currently under development, that's intended to improve on 4G. 5G promises significantly faster data rates, higher connection density, much lower latency, among other improvements.

Some of the plans for 5G include device-to-device communication, better battery consumption, and improved overall wireless coverage. The max speed of 5G is aimed at being as fast as 35.46 Gbps, which is over 35 times faster than 4G.

Key technologies to look out for: Massive MIMO, Millimeter Wave Mobile Communications etc. Massive MIMO, millimeter wave, small cells, Li-Fi all the new technologies from the previous decade could be used to give 10Gb/s to a user, with an unseen low latency, and allow connections for at least 100 billion devices . Different estimations have been made for the date of commercial introduction of 5G networks. Next Generation Mobile Networks Alliance feel that 5G should be rolled out by 2020 to meet business and consumer demands.