Chapter 3

The Physical Layer

Bandwidth (Communication): The range of frequencies transmitted without being strongly attenuated. (Measured in Hz)

Bandwidth (Computer Networks): Refers to the data rate of a channel/media i.e. the channel capacity. (Measured in bits per second)

3.1 Transmission Media

As mentioned earlier, one of the major component of any computer network is the media through which the raw bits are transmitted. All the media can be broadly classified into *Guided* or *Unguided* types. The guided media types have physical existence, while the unguided types are wireless.

3.1.1 Guided Transmission Media

Guided media types can be further categorised into two types:

1. *Point-to-Point:*
    - Provides a direct link between two devices.
    - These two devices are the only ones sharing the media.
2. *Multi-point:*
    - Media is shared.

3.1.2 Few Examples of Guided Media

1. *Coaxial Cable:*
    - Earliest LANs mostly used coaxial cables.
    - There are two types: *Thick* and *Thin* coaxial cables.
    - Thick coaxial cables uses *Vampire Taps* to connect the nodes, while the thin coaxial cables uses *BNC Connectors* for the same.
    - Since a single coaxial cable runs throughout the LAN, layout becomes simple. However, due to the shared media, only one node can transmit at any given point of time, in the entire LAN.
    - Typically provides 1GHz bandwidth.
    - Good signal-to-noise ratio.
    - Can run for several kilometers.
    - Now-a-days, mostly used for Cable TVs.

- The cross-section view of a typical co-axial cable is shown in figure 3.1

2. *Twisted Pair Cable:*
    - A very popular cable type for most LANs.
    - There are two types: *Unshielded Twisted Pair (UTP)* and *Shielded Twisted Pair (STP)*.
    - UTP cables (e.g. CAT5, CAT6) are mostly used in homes and offices, while the STP ones are used in noisy environments like factories and industries.
    - UTP consists of *4 Pairs* of wires (total 8 wires).
    - CAT5 UTP cables has a typical bandwidth of 100MHz, while the CAT6 runs at 250MHz.
    - The maximum cable segment length is *100 Meters*.
    - An UTP cable is shown in figure 3.2

3. *Optical Fiber Cable:*
    - The fiber optics, or the optical fiber cable provides very high bandwidth.
    - Unlike in the co-axial cables or in the twisted-pair cables – which uses electricity to transfer the raw bits, the data bits in a optical fiber cable is sent using *light*.
    - Individual fiber cores are of the size equal to or thinner than a human hair.
    - There are two types: *Multi-mode* and *Single-mode*.
    - The core diameter of a multi-mode fibre is *50 microns*, while a single-mode fiber's core diameter is only around *8 to 10 microns*.
    - A typical fiber optic cable is shown in figure 3.3
    - In case of a multi-mode fiber, the light travels through the fiber cable due to total internal reflection of light. (refer to figure 3.4)
    - However, in case of a single-mode fiber, since the core diameter is very small, only a single wavelength of light is guided by the fiber cable.
    - Multi-mode fibres are now-a-days mostly used for small distances (e.g. as patch cables), while single-mode fibers are used for very long distances.
    - Typically, each fiber core are used in *unidirectional* mode. As such, a pair of them are required for a full-duplex communication. So, a typical fiber cable consists of atleast two fiber cores.
    - Optical fiber cables can run for very long distances. Current single-mode fibers can have 50 Gbps bandwidth, and can run for 100 Km.
    - The infrared light used in a fiber optic cable gets attenuated to different degrees, based on its wavelength (as shown in figure 3.5). As such, the optical fiber cables are designed to operate in the 0.85 microns, 1.30 microns, and 1.55 microns, due to the lowest attenuation in those wavelength bands.

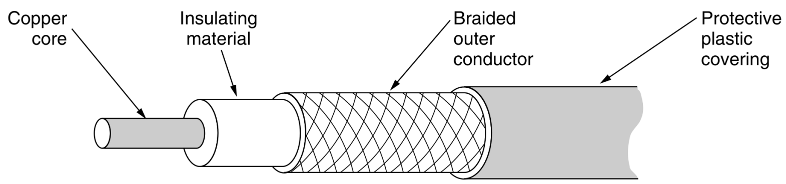


Figure 3.1: The cross-section view of a Coaxial Cable (Courtesy: Computer Networks by Tanenbaum, 5th Ed.)

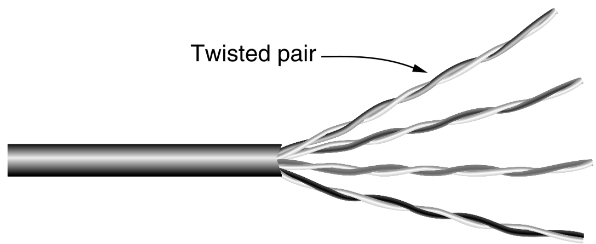


Figure 3.2: An UTP Cable (Courtesy: Computer Networks by Tanenbaum, 5^{th} Ed.)

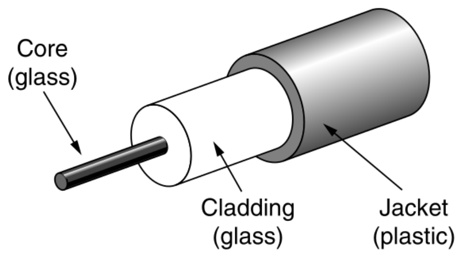


Figure 3.3: A Fiber Optic Cable (Courtesy: Computer Networks by Tanenbaum, 5^{th} Ed.)

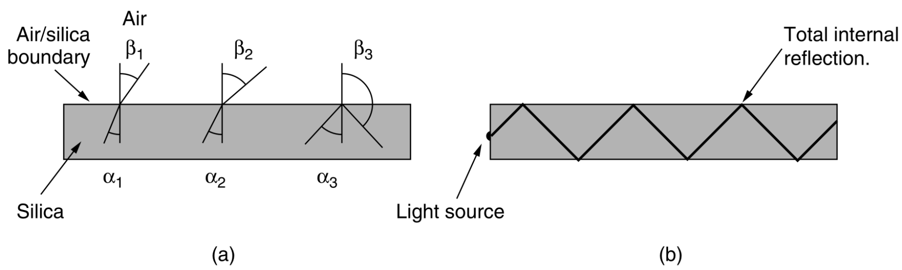


Figure 3.4: Total internal reflection in optical fiber (Courtesy: Computer Networks by Tanenbaum, 5^{th} Ed.)

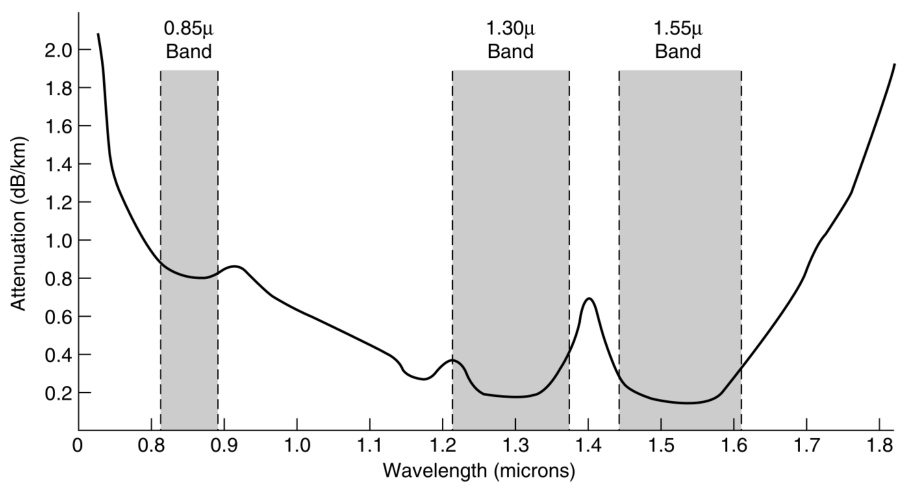


Figure 3.5: Attenuation of Infrared light in optical fiber (Courtesy: Computer Networks by Tanenbaum, 5^{th} Ed.)

3.1.3 Comparison of Fiber vs. Copper

Advantages of Fiber:

- Higher bandwidth compared to copper.
- Can run for longer distance.
- Due to the use of light for communication, optical fiber is not affected by power lines, EMI etc.
- Fiber can be laid out in dangerous locations like explosive factories, gas factories etc. (there is no risk of spark etc, due to absence of electric power in the fiber cable).
- Fiber are thinner and lightweight.
- Fiber is difficult to tap — safe from wiretappers.

Disadvantages of Fiber:

- Fiber requires skilled personnels to install.
- Can be damaged by sharp bending.
- Two-way communication requires two fibers or two wavelength bands.
- Fiber is comparatively costlier than copper.

3.1.4 Wireless Transmission

Although guided media such as optical fibre cable, and UTP cable provide high speed connectivity, these are not useful for mobile users — e.g. people going out for jogging, travelling in a car, train, aeroplane etc. For them, unguided media, or wireless transmission is the only viable option.

Electromagnetic waves can travel through space, and even in vacuum. Using appropriate sized antennas, electrical circuits can be designed which can efficiently transmit and receive these electromagnetic waves among multiple devices, which are placed some distances away, and which does not have any physical (wire) connection between them. This is the basis of wireless transmission.

Various frequency bands are used for different types of communication (both wired and wireless), as shown in figure 3.6.

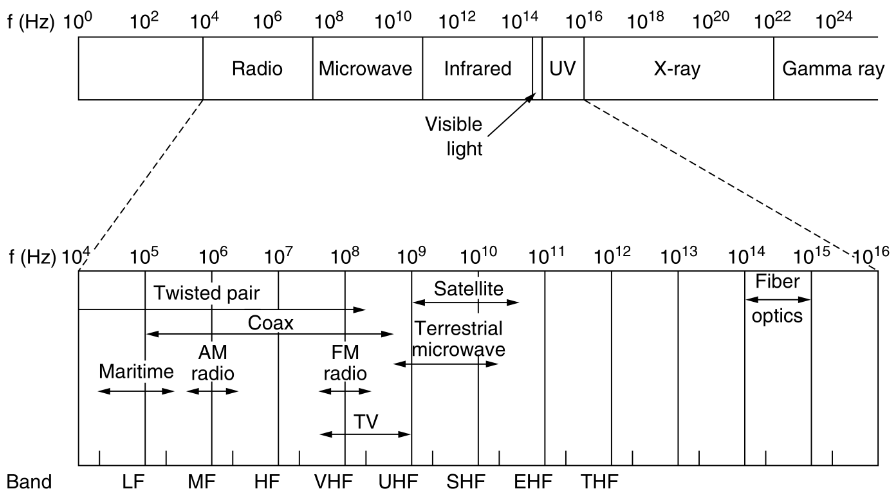


Figure 3.6: Use of electromagnetic spectrum in communication (Courtesy: Computer Networks by Tanenbaum, 5^{th} Ed.)

Frequency and wavelength of electromagnetic waves are inversely related, i.e. lower frequency has larger wavelength, and as the frequency increases wavelength decreases. Given a frequency, we can find its wavelength using the following relation:

$$\lambda \cdot f = c$$

Where, c is the speed of light (in vacuum it is $3 \times 10^8 m/sec$)[1], f is frequency in Hz, and λ is the wavelength in meters.

As such, a 300 KHz signal would have a wavelength of 1,000 meters or 1 Km. A 1 MHz signal would have a wavelength of 300 meters, 100 MHz would be 3 meters, and 1 GHz signal would have a wavelength of 0.3 meters.

Lower frequencies tend to be omni-directional in nature (useful for radio and TV broadcasts), while higher frequencies can be more directional (can be used for point to point wireless communication). Communication using higher frequencies require line-of-sight, as they tend to travel in straight lines, and they can't penetrate solid objects like buildings. However, compared to lower frequencies, higher frequencies can carry more information.

Optimal antenna size is also dependent on the wavelength (and in turn frequency) of the signal. Bigger wavelengths requires bigger antennas, while smaller wavelengths requires smaller antennas. If we look at a medium wave (MW) or short wave (SW) radio transmitter antenna, we can see its large size due to the lower frequency of transmission (and in turn bigger wavelength) used.

3.1.5 Few Examples of Wireless Transmission

1. *Microwave Transmission:*

   - Due to its very high frequency, microwaves travels in straight lines.
   - However, they can not penetrate through solid objects — as a result, a line of sight communication is required. Due to this reason, microwave antennas are generally placed on top of hills or high-rise buildings, so that the antennas are visible to each other.
   - Signals with frequencies above 4 GHz are heavily absorbed by rain (water). So, the quality of microwave transmission suffers in bad weather.

2. *Communication Satellites:*

   - Artificial satellites can be placed in orbit at different altitudes above the surface of the earth.
   - These satellites can be equipped with the required electronic devices along with the antennas, which can then be used to receive and transmit signals from the earth stations (or satellite ground stations), which can be located large distances apart.
   - Solar panels can be used to power the electronics of the satellites.
   - To avoid interference between the uplink and the downlink signals, different frequencies are assigned to them. This is handled by the satellite's *transponder*. A transponder is designed to handle multiple streams of uplink and downlink frequencies, so that multiple parties can communicate at the same time (or multiple TV channels can broadcast at the same time) using a single communication satellite.
   - Depending upon the height of the satellite from the earth's surface, multiple number of them may be required to cover a given location on the earth on continuous basis.
   - For example, if a satellite is placed at an altitude of 35,786 km from the earth's surface, it can have the same relative rotational speed as that of the earth. As a result, such a satellite would always seem to stay in the same spot in the sky, when viewed from the earth. This kind of satellites are therefore called a *Geosynchronous Equatorial Orbit* or *Geostationary Earth Orbit* (GEO) satellites.
   - An entire country's area can be covered by a single GEO satellite. To communicate across the globe, only three of them would be required.

[1] In vacuum, electromagnetic waves travels at the speed of light. However, in copper or fiber, their speed slows to around 2/3 of this value. Moreover, in solid conductors, the speed also varies with frequency.

- However, due to the high altitude of the GEOs, their latency is also quite high.
- To reduce the latency, satellites can be placed at lower altitudes. Though this reduces the latency (which can now be used for real-time communication), more number of them would be required now, since their relative orbital speed does not match with earth, and their coverage area is also small. This is shown in figure 3.7

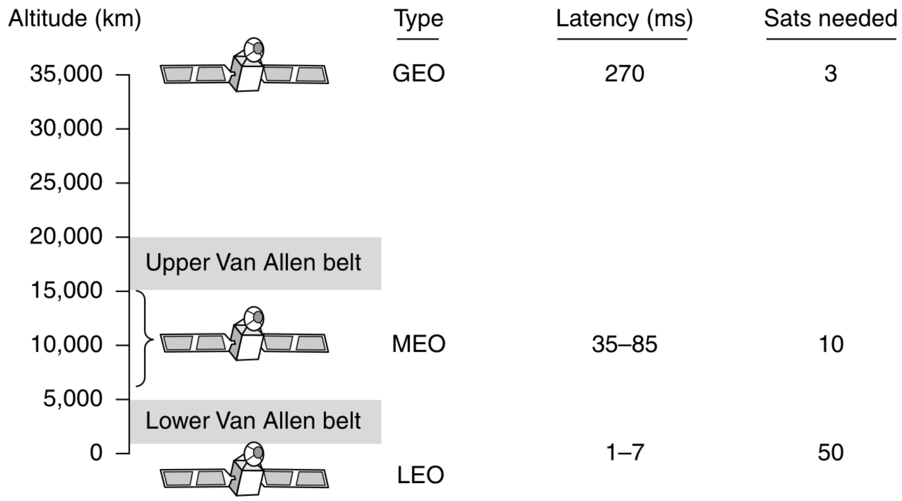


Figure 3.7: Communication Satellites (Courtesy: Computer Networks by Tanenbaum, 5th Ed.)

3.1.6 Comparison of Satellite Communication vs. Fiber

- Most of the fiber bandwidth is not directly available to the end users.
- High bandwidth of fiber is of no use for mobile users.
- Satellites can be used for easily broadcasting over a large area.
- Satellites are useful in hostile terrains (e.g. war zones), and areas with poor infrastructure.
- Satellites are useful where right-of-way is not available, or it is costly to lay fiber (e.g. remote places with low population, difficult terrain etc.)
- Satellites can be deployed rapidly — e.g. in the time of war.

3.2 Data Encoding Techniques

3.2.1 Analog vs. Digital Signal

An analog signal consists of continuous values of some sort. For example, an audio signal which represents some sound wave, may be represented in analog form with a continually varying quantity (e.g. electrical voltages) and may have a very complex waveform.

In contrast, digital signals consists of some discrete values. Digital signals can be represented in the binary form, where each discrete values can be represented using only two numbers or quantities. As a result, storing, processing, and transmitting digital signals can be much efficient, less error prone, and almost free of accumulated noise.

Most of our real-world quantities are analog in nature — e.g. sound, light, temperature etc. Due to the sampling of quantized values, a digital signal can never represent an analog quantity in exact manner. However, once we obtain the digital signal, it can be stored and transmitted without any further loss of information. This makes digital signals a preferred choice for most of the modern applications.

3.2.2 Benefits of Digital Communication

- Reliability
- Error detection & correction
- Encryption
- Compression

3.2.3 Categories of Data Encoding Techniques

Since data may be represented in both analog or digital form, we can have the following four categories:

1. Analog Encoding of Digital Information
2. Digital Encoding of Analog Information
3. Digital Encoding of Digital Information
4. Analog Encoding of Analog Information

3.2.4 Analog Encoding of Digital Information

A typical digital signal would have two variations (or two logic levels), generally referred to as '0' and '1'. These two variations may be represented by any of the constituent factors (i.e. amplitude, frequency, and phase) of an analog signal. As such, we can have the following three techniques:

(i) *Amplitude Shift Keying (ASK)*

(ii) *Frequency Shift Keying (FSK)*

(iii) *Phase Shift Keying (PSK)*

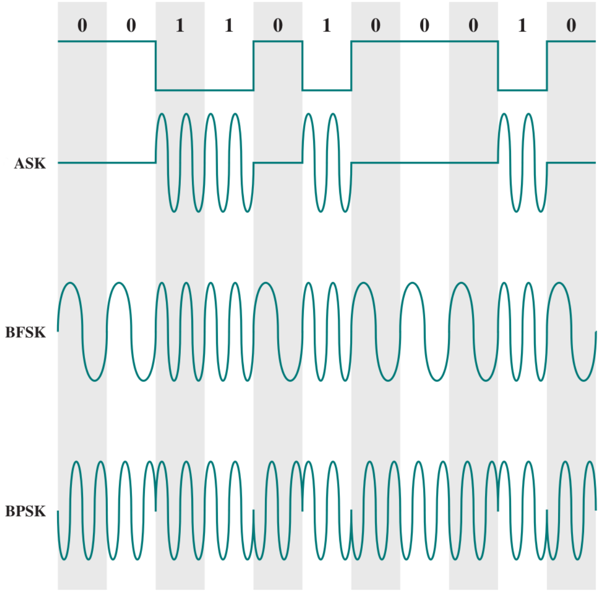
In ASK, we use two different amplitude values to represent the two digital input values, keeping the frequency and the phase constant. Similarly, in FSK[2], we use two different frequencies to represent the two digital input values, keeping the amplitude and the phase constant. And in PSK[3], the phase of the resultant signal changes, whenever there is a change in value in the input signal (i.e. the raw digital signal, which is to be encoded) from either '0' to '1', or from '1' to '0'. In PSK, the amplitude and the frequency of the signal is kept constant. The example waveforms of ASK, FSK, and PSK for a given digital input signal is shown in figure 3.8.

It should be noted that in the *Differential PSK (DPSK)* method, the phase of the resulting signal changes, whenever it encounters a '1' in the input signal (i.e. for every '1'). There is no phase change in the resulting signal for zeros in the input signal. This is shown in figure 3.8

3.2.5 Digital Encoding of Analog Information

- One way to convert an analog signal to digital representation is through *Pulse Code Modulation (PCM).*
- In this method, at uniform regular intervals a sample is taken from the input analog signal (called *sampling* or *discretization*). These discrete samples are then assigned some values from within a fixed range of values (called *quantization*).
- Though because of the discrete samples, a digital signal can not exactly represent an analog signal, for high quality reproduction we can use the *Nyquist–Shannon sampling theorem.* This theorem states that if the maximum frequency of the analog signal is *B hertz*, than taking *2B samples per second* would enable us to represent the input analog signal in a high quality digital signal.
- For example, voice data in a telephone system is limited to a maximum of 4 KHz.

[2]Or Binary FSK (BFSK), as referred to by Stallings.
[3]Or Binary PSK (BPSK), as referred to by Stallings.



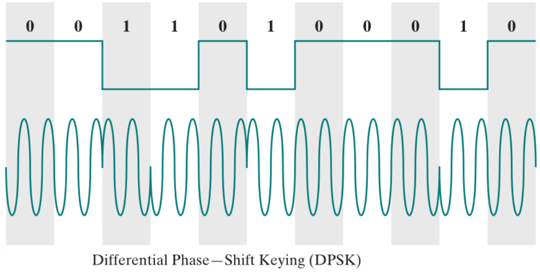


Figure 3.8: Representation of ASK, FSK, PSK, and DPSK (Courtesy: Data and Computer Communications by Stallings, 10^{th} Ed.)

- So, *8000 samples/second* is needed to faithfully represent this information.
- Each of these samples are assigned a binary code.
- Typically *8 bits* are used, which gives use *256 levels*. So, we can assign values (code) ranging from *-128* to *+127* to the samples.

3.2.6 Digital Encoding of Digital Information

Inside a computer the binary bit '1' may be represented using a positive voltage (e.g +5V), and the bit '0' may be represented using zero voltage. However, this representation is not very useful for transmitting the bits over a long distance. As such, they are converted to some other digital form before transmitting.

Non-Return-to-Zero (NRZ):

One class of such digital encoding method is known as *Non-Return-to-Zero (NRZ)*. In this method, during the transmission of a bit period, the signal does not change, or it does not go from one voltage level to another, crossing the zero level.

There are various types of NRZ schemes. Some of them uses *unipolar* (i.e. +ve and 0V), while others uses a *bipolar* (i.e. +ve and -ve) voltage. The bipolar scheme has some advantages compared to the unipolar ones. First, because the peak-to-peak voltage level can be higher, they can be transmitted for longer distances. Second, since they are essentially AC signals now, bipolar signals can pass through existing cabling infrastructure which might use isolation transformers, impedance matching transformers, or coupling/decoupling capacitors etc.

Two examples of NRZ method are:

(i) *NRZ-L (Non-Return-to-Zero Level):* Different implementations may use different voltage levels to depict the NRZ-L method. However, for showing an example, we may use *RS-232*, which is used for serial communication and is an industry standard. The original PCs used a RS-232 port (or serial port) to connect modems, mouse, printers etc. The RS-232 uses a bipolar voltage level. The '1' is represented by -ve (typically -12V to -5V), and '0' is represented by +ve (+5V to +12V). This is shown in figure 3.9

(ii) *NRZ-I (Non-Return-to-Zero Inverted):* In this method also, voltage level used to represent a particular bit is dependent on the specific implementation. For example, in *RLL Codes*, absence of transition represents bit '0', while transition from either *Low-to-High* or *High-to-Low* represents a '1'. This is shown in figure 3.9. However, *HDLC* and *USB* uses the opposite convention — i.e. the '1' is transmitted without any transition, while for each '0', a transition is used.

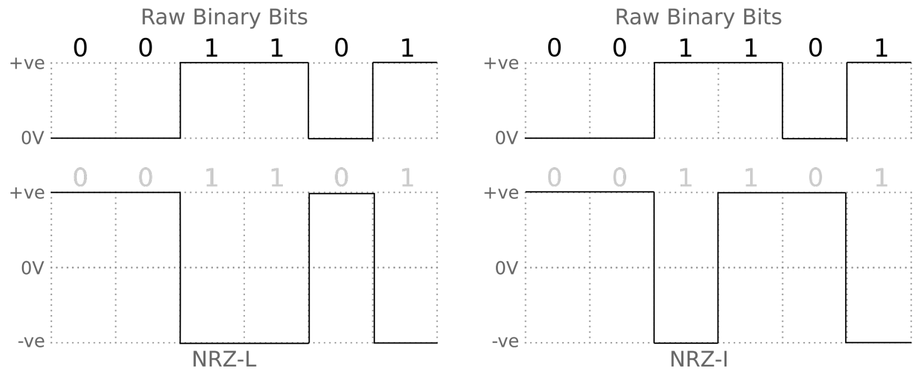


Figure 3.9: Waveforms for NRZ-L and NRZ-I Encodings

The NRZ-I belongs to a class called *differential encoding* method. In this type, the bits are not represented with absolute voltage values, but is dependent on the previous bit state. As a result, even if the transmit (TX) and receive (RX) wires are interchanged either in the source or in the destination device, the overall meaning of the bits are unaltered — i.e. the receiver is still able to detect the 1s and 0s correctly. This is not the case in regular encoding methods like in NRZ-L. The receiver would detect a '1' as '0' and a '0' and '1', if the TX and RX wires are interchanged.

Disadvantages of NRZ methods:

- Hard to tell when one bit ends or starts.
- With long strings of 0s or 1s, timing drift between the sender and the receiver may result in error.

Manchester Encoding:

The Manchester encoding is used by the earlier versions of the *Ethernet*. It uses bipolar encoding, thereby eliminating the DC component. In this method, there is always a transition (either from Low-to-High or from High-to-Low) at the middle of each bit period, as shown in figure 3.10. The advantage of this transitions are:

- *Self Clocking:* Since each bit will always have a transition at the middle, the receiver can always synchronise its clock correctly, thus avoiding clock drifting errors present in the NRZ schemes.
- *Well Balanced:* Since each bit will always have both the positive and the negative component, even if long strings of 0s or 1s are present in the input, it can't make the line/cable more positive or more negative, but keeps the line in a neutral state. This eliminates interference with the nearby cables.

How the bits '1' and '0' are represented in Manchester encoding is a little bit confusing. Many authors, including Tanenbaum specifies the transition for '1' as from 'High-to-Low', and 'Low-to-High' for '0'. However, authors like Stallings uses the opposite convention. The convention used by Stallings is based on the IEEE 802.4 (Token Bus) standard, as well as the lower speed versions of the IEEE 802.3 (Ethernet) standard.

The Ethernet and the Token Bus standards represents the bit '1' with a *Low-to-High* transition, and the bit '0' with a *High-to-Low* transition. This convention is shown in figure 3.10.

One disadvantage of Manchester encoding is that, compared to other schemes like NRZ, it requires double the clock frequency, since each bit does a transition at the middle.

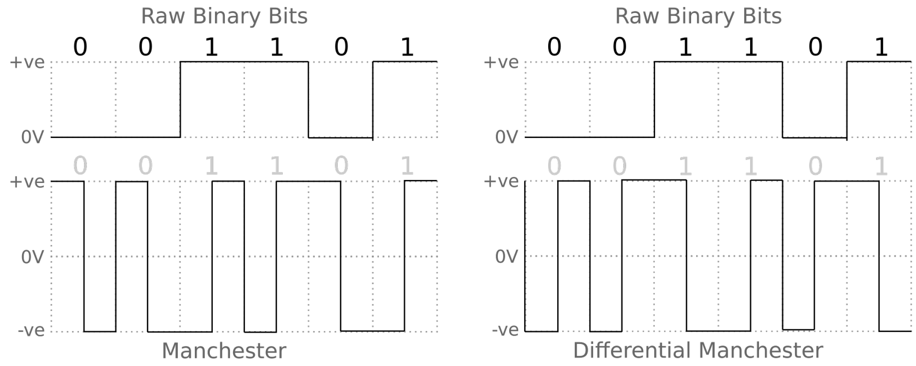


Figure 3.10: Waveforms for Manchester and Differential Manchester Encodings

Differential Manchester Encoding:

Differential Manchester encoding is also a bipolar encoding. Moreover, it also uses a transition at the middle of each bit period. As such, this method also includes the *self clocking* and the *well balanced* advantages. It also has the disadvantage of requiring double the clock frequency.

However, since this method uses a differential encoding method, it is immune to the wire swap problem (the regular Manchester encoding suffers from this problem).

One convention of the differential Manchester encoding uses a *transition at the beginning of a bit* to represent a '0', and *absence of transition* to represent a '1'. This is shown in figure 3.10.

Differential Manchester encoding is used in the IEEE 802.5 (Token Ring) standard.

3.2.7 Analog Encoding of Analog Information

If we wish to transmit multiple analog signals through a given media (either guided or unguided), and if they occupy similar frequency bands (for example, transmitting multiple voice signals), then there will be interference among the signals, and transmission would be unsuccessful. In this case, we would need to shift each analog signal to an independent non-overlapping frequency band.

Similarly, for efficient use of a given media's bandwidth, we may transmit multiple analog signals through it, but in independent non-overlapping frequency bands, so that there is not interference among them.

Moreover, as noted earlier, the optimal size of an antenna depends on the wavelength of the signal to be transmitted. For lower frequency signals like voice, the size of the antenna would be too big, and would be infeasible to build. So the signal would need to be shifted to some higher frequency for practical and efficient transmission.

All these problems can be solved by using a technique called *modulation*. In simple terms, modulation involves combining the input analog signal (called the *modulating signal*) with some other higher fixed frequency signal (called the *carrier signal*). The resulting signal can then be transmitted through some media, and at the other end (i.e. the receiver), the carrier signal is subtracted from the received signal to recover the modulating signal (called *demodulation*).

Types of Modulation:

(i) *Amplitude Modulation (AM)*

(ii) *Frequency Modulation (FM)*

(iii) *Phase Modulation (PM)*

In all the modulation techniques, some characteristics of the carrier signal is modified in accordance with the modulating signal (i.e. the signal intended to be transmitted). In case of AM, the amplitude of the carrier signal changes in proportion to the amplitude of the modulating signal. The frequency and the phase of the carrier signal remains the same. In case of FM, the frequency of the carrier signal changes in proportion to the amplitude of the modulating signal. The amplitude and the phase of the carrier signal remains the same. The resulting waveforms for AM, PM, and FM for a given modulating signal (a sine wave) is shown in figure 3.11.

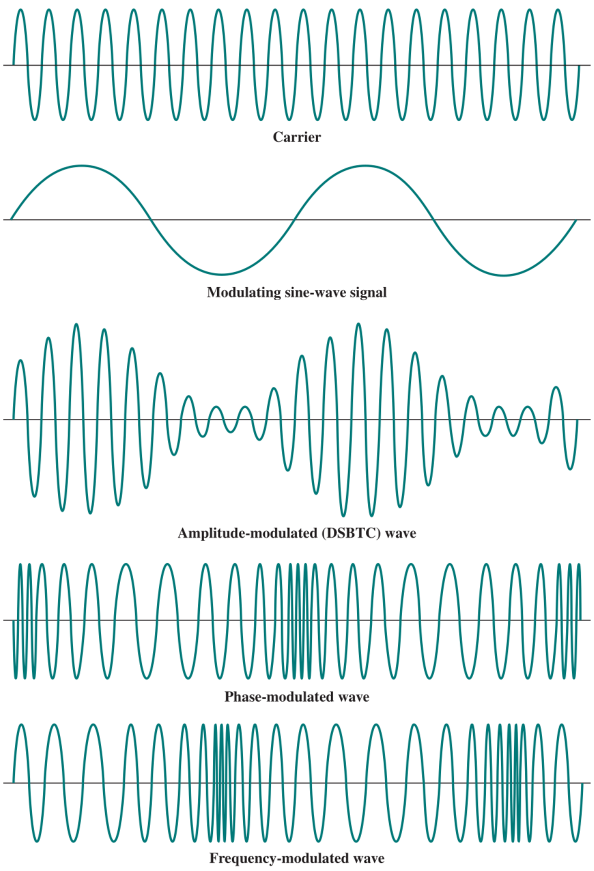


Figure 3.11: Waveforms for Amplitude Modulation, Phase Modulation, and Frequency Modulation (Courtesy: Data and Computer Communications by Stallings, 10^{th} Ed.)

3.3 Multiplexing

- Under the simplest conditions, a media can carry only one signal at any moment of time.
- For multiple signals to share one media, the media must be somehow divided giving each signal a portion of the total bandwidth.
- Transmission services are very expensive.
- Multiplexing and compression techniques save business money.
- As data capacity of line increases, it becomes more cost effective for the company to send multiple signals through a single line.
- Most data devices requires modest data rate support — so a single link can be shared to send multiple signals without effecting the end user.
- A logical block diagram of a multiplexer is shown in figure 3.12, wherein n inputs are combined by the multiplexer and sent through a single link. At the other end, a demultiplexer is used to separate the signal into n outputs again.
- A common use case of a multiplexer is in cable TV transmission. Here, instead of laying multiple cables to each of the households to transmit multiple TV channels (which would be very costly,

and difficult to manage as well), all the required TV channels are multiplexed and sent through a single high-bandwidth co-axial cable.

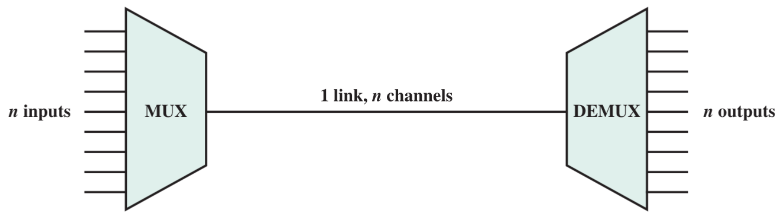


Figure 3.12: Multiplexer (Courtesy: Data and Computer Communications by Stallings, 10^{th} Ed.)

3.3.1 Alternative Approaches to Multiplexing

If multiplexing is not used, we may use the following alternative methods. However, they might not be suitable for certain requirements.

- *Direct Point-to-Point Links:*
  - In this case, each end device (node) is connected to the source using an independent cable.
  - This will require I/O ports for each of the target devices.
  - This will also need large amount of wiring, more so if the devices are located on different floors etc.
- *Multi-drop Link:*
  - In this case, all the nodes are connected to a single cable.
  - The host polls the clients to see who wants to send.
  - This method save on I/O ports, and the amount of cabling required.
  - However, since a single cable is shared by all the nodes, total communication load is not greater than the data rate of the cable.

3.3.2 Multiplexing Techniques

- *Frequency Division Multiplexing (FDM)*
- *Wavelength Division Multiplexing (WDM)*
- *Time Division Multiplexing (TDM)*
- *Code Division Multiplexing (CDM)*

Frequency Division Multiplexing (FDM):

- This is the oldest multiplexing technique.
- Popular in the broadcasting of radio, television, and cable TV.
- Frequency spectrum is divided among the logical channels.
- Each user has exclusive access to his/her channel.
- Sends signals in several distinct frequency ranges (e.g. cable TV).
- Each signal is modulated onto different carrier frequency and carrier frequencies are separated by guard bands, as shown in figure 3.13

- Bandwidth of the transmission media must exceed the required bandwidth of all the constituent signals.
- Analog signalling is used to transmit the signals.
- Non-overlapping frequency ranges are assigned to each 'user' or signal on a media. Thus all the signals are transmitted at the same time, each using different frequencies.
- A multiplexer accepts the inputs and assigns frequencies to each device.
- The multiplexer is attached to a high-speed communication link.
- A corresponding demultiplexer is located on the other end of the high-speed line and separates the multiplexed signals.
- Since it involves analog signalling, it is more susceptible to noise.
- FDM can't utilise the full capacity of the cable.
- There must be considerable gap between the frequencies bands in order to ensure that signals from one band do not affect signals in another band.
- FDM is mostly used to carry analog signals. Though modulated digital signals can also be sent.

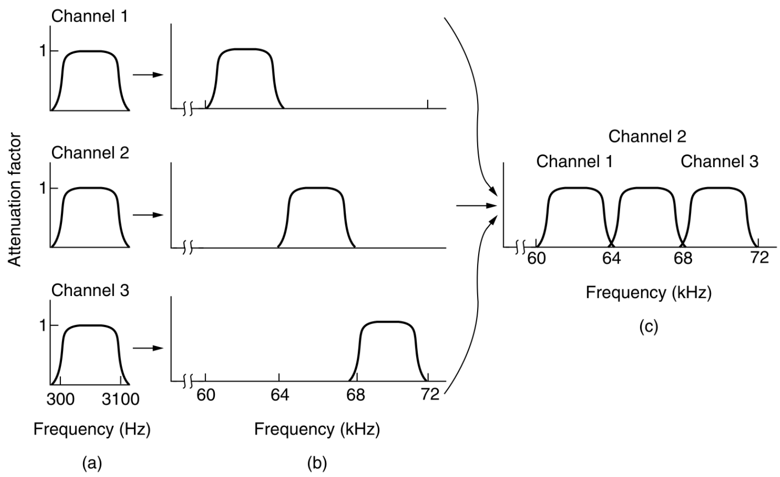


Figure 3.13: Frequency Division Multiplexing (FDM) (Courtesy: Computer Networks by Tanenbaum, 5$^{\text{th}}$ Ed.)

Wavelength Division Multiplexing (WDM):

- WDM is similar in operation to FDM, but applied to fibers (light wavelengths). It is shown in figure 3.14
- Different wavelength lasers (called *lambdas*) are used to transmit multiple signals.
- Each signal carried on the fiber can be transmitted at a different rate from the other signals.
- There are two types:
  (i) *Dense WDM (DWDM):* Combines many lambdas into a single fiber.
  (ii) *Coarse WDM (CWDM):* Combines only a few.

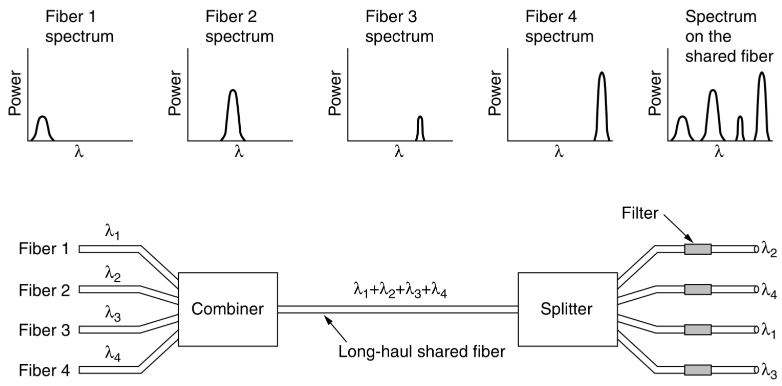


Figure 3.14: Wavelength Division Multiplexing (WDM) (Courtesy: Computer Networks by Tanenbaum, 5$^{\text{th}}$ Ed.)

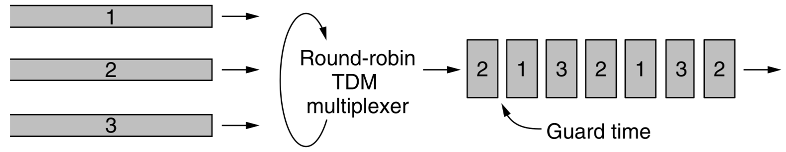


Figure 3.15: Time Division Multiplexing (TDM) (Courtesy: Computer Networks by Tanenbaum, 5$^{\text{th}}$ Ed.)

Time Division Multiplexing (TDM):

- Like FDM, TDM saves money by allowing more than one user to use a media at the same time.
- In this technique, sharing is done by dividing the available transmission time on a media among the users, as shown in figure 3.15
- Each user periodically gets the entire bandwidth for a small amount of time.
- Digital signalling is used exclusively.
- There are two types:

  (i) *Synchronous TDM:*
    - Instead of dividing the bandwidth into frequency bands, TDM splits it into time slots.
    - Each channel is given a regular time slot in which to send a signal.
    - The users take turns, each one having exclusive use of the media, in round robin fashion.
    - Data is divided into frames. Each time slot can carry one frame.
    - In Synchronous TDM, time slots are transmitted irrespective of whether a user/source has data to send or not.
    - Empty time slots degrades the utilization of the link.
    - Examples of Synchronous TDM are: T1, ISDN.

  (ii) *Statistical or Asynchronous TDM:*
    - In this technique, time division is done on demand rather than on a fixed basis.
    - Data is broken down into smaller packets.
    - The link is rescheduled on a per-packet basis, i.e. whoever has data to send initiates a request, and its packet(s) are transmitted in the next available slot.
    - This requires that each packet contains a terminal identifier, so that at the other end it can be correctly delivered to the intended user.
    - Packets from different sources are interleaved on the link.

- May require storage buffers for the packets that are contending for the link.
- For variable length data, a length field can also be included.
- Packet queue may be processed FIFO or using some other technique.
- Buffer overflow is called congestion.
- Transmits data only from the active nodes. No space is wasted for the inactive nodes.
- Allows connection of more nodes to the circuit than the capacity of the circuit.
- This works on the premise that not all nodes will transmit at full capacity at all times.

In practice, if a given link has high bandwidth, both FDM and TDM may be combined:

- The available bandwidth is broken up into multiple frequency bands (FDM).
- In each band multiple channel can be accommodated using TDM.

3.4 Switching

In a network, if the sender and the receiver is not directly connected to each other with a direct link, but needs to go through some intermediate nodes, we have to find a way to send the communication from the source to the destination through these intermediate nodes. This involves some kind of switching (connection establishment) among the nodes.

3.4.1 Types of Switching

(i) *Circuit Switching:*

- This is a typical scenario in an old PSTN. Here, before the two users can communicate, the intermediate telephone exchanges has to hunt for an available outgoing link to the next exchange. Once all the exchanges and the local loop (link to the end user from the nearest exchange) has been connected (switched), the conversation can start. This is shown in figure 3.16
- In circuit switching, an end-to-end physical path (called the *circuit*) is established *before* any data can be sent.
- Once the path is established, it is maintained till the call is finished.
- Once the path is established, the only delay for data is the propagation delay.
- No congestion problem can occur, once the path is setup, since the entire physical path exclusively used by the two users only.
- Busy signal may be received *before* the path is established, due to lack of switching or trunk capacity.

(ii) *Message Switching:*

- No physical path is established in advance between the sender and the receiver.
- When a sender has a block of data to send, it is stored in the first switching office (in this case, a router), and then forwarded later, one hop at a time.
- That is, the intermediate router after storing the data, would look for an outgoing link to the next router. When one becomes available, the two adjacent routers establishes a connection, the data is transferred, and the connection is terminated. This process goes on, till the data is received by the destination node.
- In this method, no end-to-end physical connection is made. Instead, at a time, connections are temporarily established between two nodes only.
- Each block of data is received in its entirety, inspected for errors, and then retransmitted. A network using this technique is called a *store-and-forward* network.

(iii) *Packet Switching:*

- With message switching, there is no limit on the block size.
  - This means the intermediate routers must have disks to store large blocks.
  - Large blocks can tie-up a router-to-router link for minutes — making interactive traffic infeasible.
- Packet switching networks places a tight upper limit on the block size, allowing blocks to be buffered in router's main memory, instead of disk.
- This also means that a single user can not tie-up a transmission line for a long time.
- Moreover, a packet of a multi-packet message can be forwarded before the next packet has fully arrived — this improves throughput and reduces delay.
- The operational diagram of circuit switching and packet switching is shown in figure 3.17

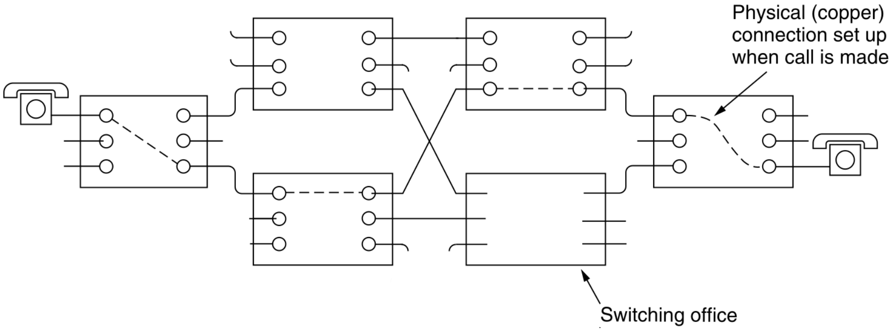


Figure 3.16: Setting up a circuit in a Public Switched Telephone Network (PSTN) (Courtesy: Computer Networks by Tanenbaum, 5$^{\text{th}}$ Ed.)

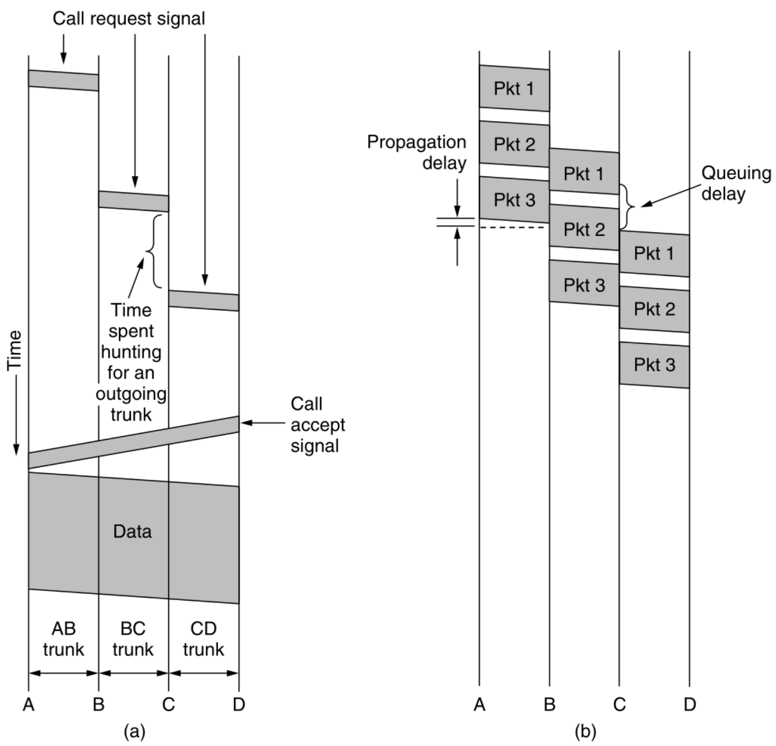


Figure 3.17: Circuit Switching & Packet Switching (Courtesy: Computer Networks by Tanenbaum, 5$^{\text{th}}$ Ed.)

3.4.2 Difference between Circuit Switching & Packet Switching

- Circuit switching needs advance path setup.
- Circuit switching reserves the bandwidth all the way up from the sender to the receiver. So all the data bits follows the same path — as a result, data can not arrive out-of-order.
- In packet switching, packets may arrive out-of-order, if they follow different paths.
- Packet switching is more fault tolerant than circuit switching.
- In packet switching, packets may have to wait before they are forwarded – since no bandwidth is reserved in advance.
- Congestion may occur at different times. In circuit switching, it may occur at circuit setup time, while in packet switching, it may occur when the packets are sent.
- Bandwidth may be wasted in case of circuit switching, if the user has nothing to send for some time. There is a trade-off between guaranteed service & wasting resources vs. not guaranteeing service & not wasting resources.
- In case of packet switching, the store-and-forward technique adds delay.
- Circuit switching is transparent to the user — the sender and the receiver can use any bit rate, format, or framing method. But in packet switching, the carrier determines the basic parameters.

It should be noted that almost all of our modern networks are packet switched. Circuit switching may be used in case of dial-up connection and such. However, message switching is never used.