Computer Networks

Module 1: Introduction

Module 2: Physical Layer

Physical layer is the bottom-most layer in the Open-System Interconnection (OSI) model which is a physical and electrical representation of the system.

It consists of various network components such as power plugs, conductors, receivers, cable types, etc.

The physical layer sends data bits from one device to another device.

It is responsible for the communication of the unstructured raw data streams over a physical medium.

1. Types of Media
2. Guided Media

In guided media, the signals are guided by a physical medium of transmission. Signals are under control and remains in a physical wire. E.g. Copper wire, optical fiber.

* Copper:

1. **Coaxial Cable** – Consists of an inner conductor and an outer conductor separated by an insulator. The inner conductor is usually copper. The outer conductor is covered by a plastic jacket.

Typical diameter: 0.4 inches to 1 inch

Application: TV cables

Advantages: High bandwidth, attenuation is less.

1. **Twisted Cable –** Consists of several insulated copper wires (generally 1mm thick) twisted together in a helical form. The purpose of twisting the cables is to reduce cross-talk interference between several pairs of wires.

Advantages: Much cheaper than coaxial cables

Disadvantages: Susceptible to noise and electromagnetic interference, Attenuation is large.

* Unshielded twisted pair – No insulation is provided. Hence more susceptible to noise and interference.
* Shielded twisted pair – A protective thick insulation is provided. Its expensive and not generally used.
* Optical Fiber

In optical fiber, light is used to send data. In general terms, presence of light is taken as bit 1 and absence is taken as bit 0.

Optical fiber consists of inner core of generally glass or plastic.

The core is surrounded by a cladding of the same material but of different refractive index.

This cladding is surrounded by a plastic jacket which prevents the optical fiber from EM interference and harshly environment.

Advantage: Better bandwidth, almost no attenuation, very low EM interference.

Disadvantage: Endpoints of optical wires are fairly expensive.

1. Based on material

* Glass optical fiber
* Plastic optical fiber

1. Based Radius
   * Thin optical fiber
   * Thick optical fiber
2. Based on light source
   * LED (Light Emitting Diode) – For low bandwidth
   * IBD (Injection Based Diode) – For high bandwidth
3. Unguided Media [Wireless transmission]

Unguided media means that there is no physical path for the signal to propagate. These are basically electromagnetic waves. E.g. radio waves.

1. Radio – Radio is a general term for any kind of frequency. However, generally lower frequencies are termed as **radio waves** and higher frequencies are termed as **microwaves.**

Applications include wireless keyboard, LAN, wireless Ethernet.

Depending on frequency, Radio offers different bandwidths

1. Terrestrial Microwave – Here two antennas are used for signal propagation. One antenna serves as the sender and another as the receiver.
2. Satellite Communication – Satellite acts as a switch in the sky. Generally, one station on earth transmits signal to satellite and it is received by many stations on earth.
3. Transmission Mode

Transmission mode can be categorized into ->

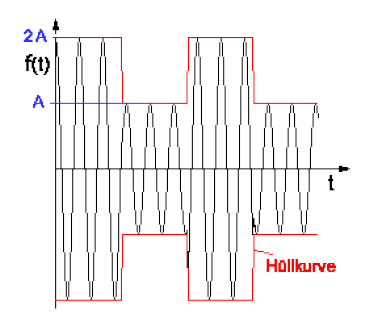
1. Simplex: Signal can be transmitted only in a single direction. E.g. TV broadcast
2. Half duplex: Signal transmission is bi-directional but it can be transmitted in only one direction at a time. E.g. walkie talkie
3. Full duplex: Signal transmission is bi-directional and it can be transmitted in both directions simultaneously. E.g. talking in telephone.
4. Communication Links

Communication links connect two nodes in a network. These can be categorized into ->

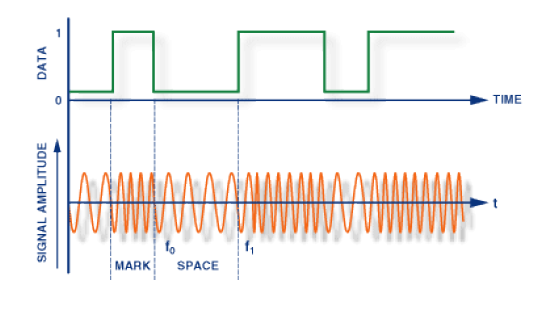
1. Point-to-point – Only two nodes are connected together. The signal sent by the sender can only be received by a single node which is on the other side.
2. Multipoint – In this communication link type, the signal can be received by more than one node.
3. Problems associated with transmission of signals
4. Attenuation: As the signal transmits long distances in a network, the quality of the signal slowly degrades. This is called attenuation. This can be managed by using amplifiers or repeaters at certain intervals in the network.

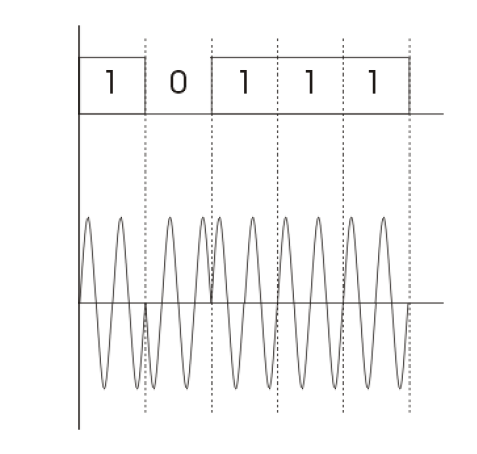
Amplifiers increase the amplitude/strength of the signals whereas repeaters regenerate the signal.

1. Noise: In a communication channel many signals get transmitted simultaneously, also certain random signals are also present in the medium. Due to interference of these signals, our signal gets disrupted a bit.
2. Data Encoding
3. Digital data to analog signals
4. Amplitude Shift Keying (ASK) – It represents digital data as variations in the amplitude of the career wave. Two different amplitudes represent 0 and 1.



1. Frequency Shift Keying (FSK) – It represents digital data as variations in the frequency of the career wave.



1. Phase Shift Keying (PSK) – Here, the phase of the career wave is shifted whenever variation in the digital data is encountered.
2. 
3. Digital data to digital signals
4. Unipolar (NRZ) – 0: Low, 1: High
5. Polar:
   * NRZ-Level – 0: High. 1: Low
   * NRZ-Inverted – 0: No transition, 1: Transition
   * RZ -
   * Biphase
     + Manchester
     + Differential Manchester
6. Bipolar
7. Multilevel
8. Multitransitional
9. Analog data to digital signal [Digitization]

Sampling frequency must be at least **twice** than the highest frequency in the signal so that it can be fairly regenerated.

1. **Pulse code modulation (PCM)** – Here intervals are equally spaced. 8bit PCM uses 256 different amplitude levels.
2. **Delta Modulation (DM**) – Since successive samples do not differ much, we just send the difference between the previous and present samples instead of sending the samples themselves. It requires fewer bits than PCM.
3. Multiplexing

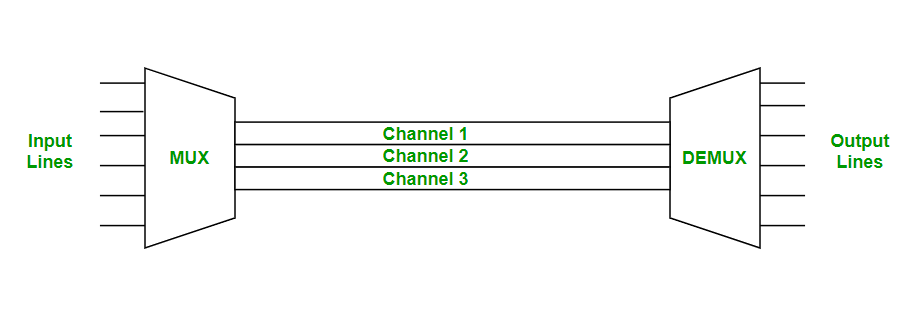
When two communicating nodes are connected through a media, it generally happens that the bandwidth of the communication medium is much higher than that of the nodes. Hence, the whole capacity of the link is not utilized in this case. The link can be further exploited to send several signals combined into one. This combining of several signals into one is called **multiplexing**.

In other words, the process in which multiple signals coming from multiple sources are combined and transmitted over a single communication line is called **Multiplexing**.

1. Frequency Division Multiplexing (FDM)

Frequency Division Multiplexing is defined as a type of multiplexing where the bandwidth of a single physical medium is divided into number of smaller, independent frequency channels.

It’s used in radio and television transmission.

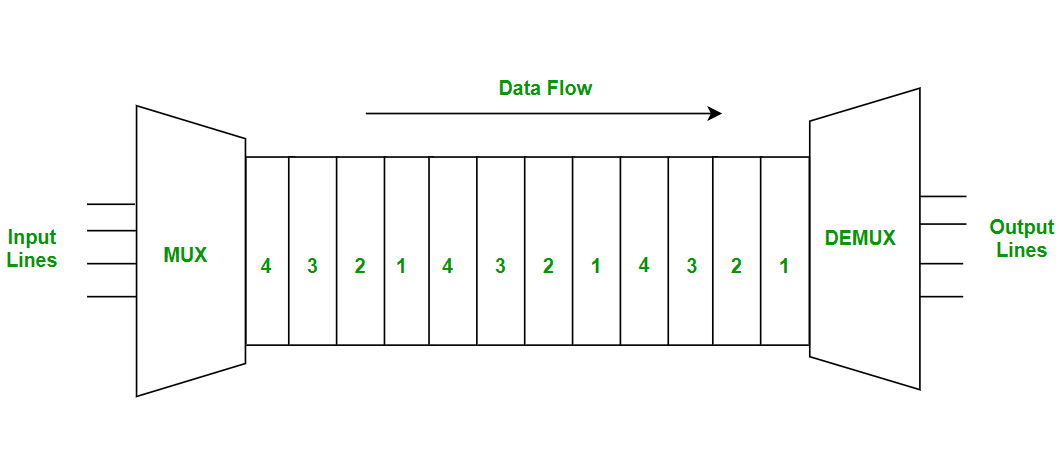


In FDM, we can observe a lot of inter-channel crosstalk interference. To avoid this, unused strips of bandwidth (guard bands) must be placed between each channel.

1. Time Division Multiplexing (TDM)

Time-division multiplexing is a method of putting multiple data streams in a single signal by separating the signal into many segments, each having a very short duration. Each individual data stream is re-assembled at the receiving end based on timing.

In TDM, all signals operate with the same frequency (bandwidth) at different times.



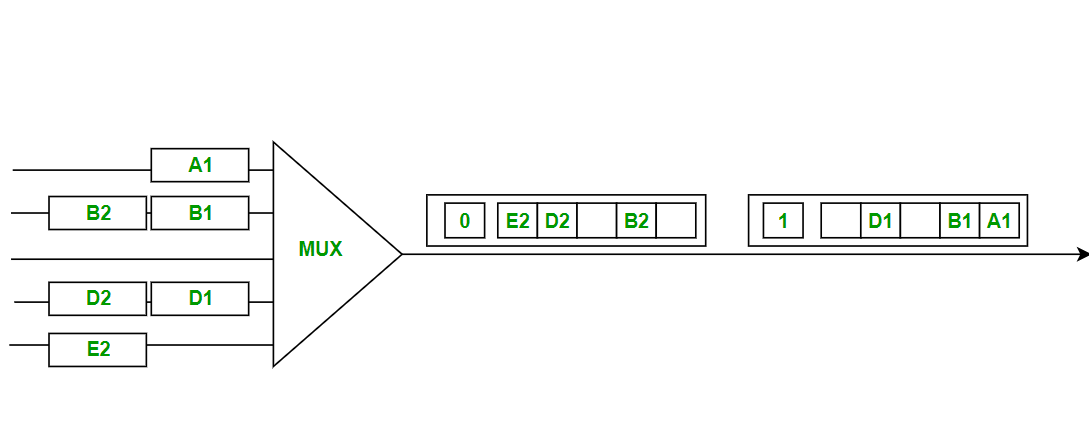
Types of TDM ->

1. **Synchronous Time Division Multiplexing**

Synchronous TDM is a type of multiplexing where the input frame already has a slot in the output frame. Time slots are grouped into frames. One frame consists of one cycle of time slots.

Synchronous TDM is not efficient because if the input frame has no data to send, a slot remains empty in the output frame.

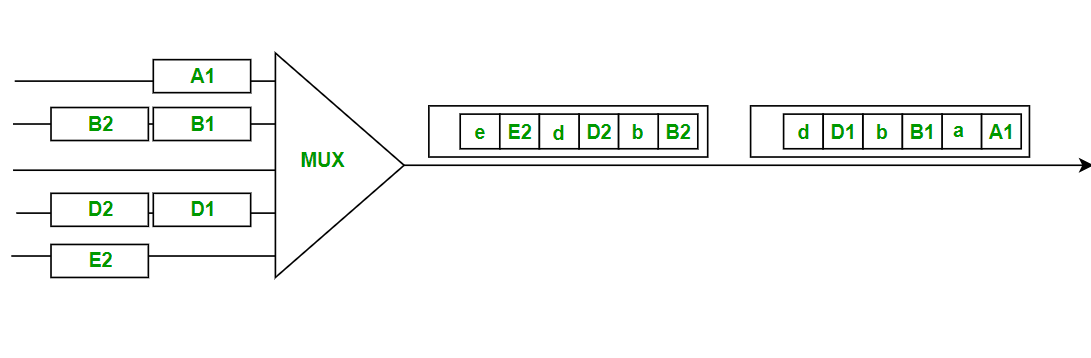
Here, we need to mention the synchronous bit at the beginning of each frame.



1. **Statistical (Asynchronous) Time Division Multiplexing**

Statistical TDM is a type of multiplexing where the output frame collects data from the input frame till it is full, not leaving an empty slot like synchronous TDM.

Here, we need to include the address of each particular data in the slot that is being sent to the output frame.



1. Wavelength Division Multiplexing (WDM)

Wavelength Division Multiplexing is a multiplexing technology used to increase the capacity of optical fiber by transmitting multiple optical signals simultaneously over a single optical fiber cable, each with different wavelength.

Each signal is carried on a different wavelength of light, and the resulting signals are combined onto a single optical fiber for transmission. At the receiving end, the signals are separated by their wavelengths, de-multiplexed and routed to their respective destinations.

WDM is an analog multiplexing technique.

Categories of WDM ->

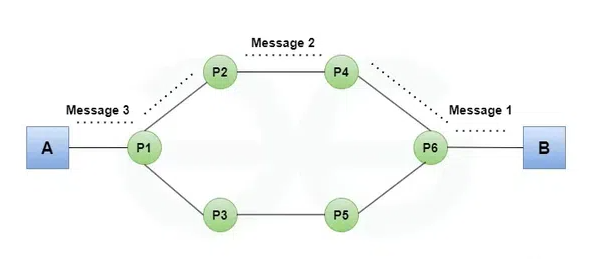
1. Dense Wavelength Division Multiplexing (DWDM) – Used to multiplex large number of optical signals to a single fiber, typically up to 80 channels with a spacing of 0.8 nm or less between the channels.
2. Coarse Wavelength Division Multiplexing (CWDM) – Use for lower capacity applications, typically 18 channels with a spacing of 20nm between the channels.

Applications of WDM include telecommunications, cable TV, internet service providers, and data centers.

1. Circuit Switching, Message Switching, Packet Switching

**Circuit Switching:** Before OSI model and TCP/IP model, telephone networks were used. To physically connect these telephone networks, circuit switching was used.

In circuit switching, network resources (bandwidth) are divided into pieces and the bit delay is constant during a connection. The dedicated path/circuit established between the sender and receiver provides a guaranteed data rate. Data can be transmitted without any delays once the circuit is established.



In circuit switching, a dedicated communication path or circuit is established between two end points for the duration of the communication session. This path remains reserved exclusively for the whole duration of the session.

Features ->

* Dedicated Path
* Fixed Bandwidth
* Connection Setup and Teardown: Establishment [Setting up the circuit] -> Data Transfer -> Teardown [Disconnecting the circuit]
* Once the circuit is setup, there is minimal delay during the communication since the path is already reserved. However, there can be significant delay during the initial setup phase.
* Resource Reservation

Advantages ->

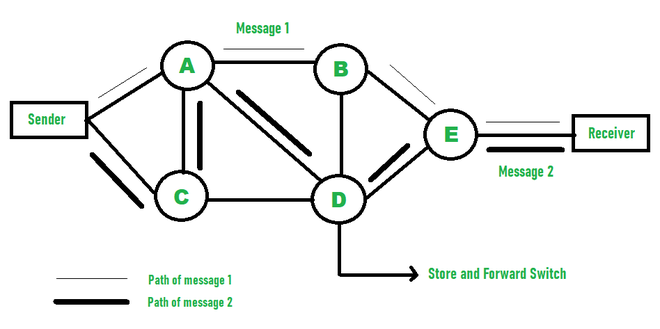
* Consistent Quality of Service [QoS]
* Low latency during communication
* Reliability for continuous data streams

Disadvantages ->

* Inefficient use of resources
* Setup time
* Scalability

**Packet Switching:** Unlike Circuit switching, packet switching divides the data into packets that are transmitted over a shared network rather than a dedicated network. Each packet may take a different path to the destination, and there is no need for establishing a dedicated path beforehand.

This method is more efficient and scalable, making it ideal for data networks like the Internet.



Packet switching uses the **Store and forward** while switching the packets -> while moving forward the packet, each hop first stores the packet and then moves forward.

Types of delays ->

* Transmission delay
* Propagation delay
* Queuing delay
* Processing delay

Advantages over Circuit Switching ->

* More efficient in terms of bandwidth (since concept of reserving a circuit is not three).
* Minimal transmission latency
* More reliable (as destination can detect missing packet)
* Cost-effective

Disadvantages ->

* Doesn’t transfer packets in a particular order unlike circuit switching
* Since the packets are unordered, we need to provide sequence numbers for each packet.
* Complexity is more.
* Transmission delay is more because of rerouting.
* Packet switching is only beneficial for small data, but for large data, CS is better.

Types of packet switching ->

* **Connection-oriented packet switching (Virtual Circuit)**:

Before starting the transmission, it establishes a logical path or virtual connection using a signaling protocol between the sender and receiver. All the packets that belong to the flow will follow this predefined route.

Virtual Circuit ID is provided by switches/routers to uniquely identify this virtual connection.

* **Connectionless Packet Switching (Datagram)**

Unlike connection-oriented PS, in connectionless PS each packet contains all necessary addressing information such as source address, port numbers, etc.

1. Spread spectrum
2. Topology

**Network topology** is the arrangement of a network that comprises if nodes and connecting lines via sender and receiver.

1. Point-to-point Topology
2. Mesh Topology
3. Star Topology
4. Bus Topology
5. Ring Topology
6. Tree Topology
7. Hybrid Topology

Module 3: Data link Layer

The data-link layer is the second layer in the OSI model, above the physical layer, which ensures that error free data is transferred between the adjacent nodes in the network.

It breaks the datagrams passed down by above layers and converts them into frames ready for transfer. This is called **framing.**

It provides two main functionality ->

* Reliable data transfer service between two peer network layers.
* Flow control mechanism which regulates the flow of frames such that data congestion is not there at slow receivers due to fast senders.

1. Data-Link Control

The data link control is responsible for transmission of message over transmission channel using techniques like framing, error control, and flow control. These are discussed below

1. Framing

Since the physical layer merely accepts and transmits a stream of bits without any regard to the meaning or structure, it is up to the data link layer to create and recognize frame boundaries. This can be accomplished by attaching special bit patterns at the beginning and end of the frame.

If these bit patterns accidentally occur within the data, special care must be taken to make sure that these patterns are not mistaken as frame delimiters.

Framing methods ->

1. Character Count – This method is rarely used and is generally required to count the total number of characters that are present in the frame.

This method uses a field in the header to specify the number of characters in a frame.

When the data link layer at the destination sees the character count, it knows how many characters follow, and hence where the end of the frame is.

Disadvantage -> If somehow the character count is accidentally modified due to some transmission error, the destination will lose synchronization and will be unable to find the start of the next frame.

1. Starting and Ending characters with **Character stuffing** – Here, each frame starts with the ASCII character sequence **DLE STX** and ends with the sequence **DLE ETX.**

DLE -> Data Link Escape

STX -> Start of TeXt

ETX -> End of TeXt

This method overcomes the drawbacks of the character count method. If the destination ever looses synchronization, all it has to do is look for DLE STX and DLE ETX character sequences.

However, if binary data is being transmitted, there exists a possibility that DLE STX and DLE ETX occurs in the data itself. Since this can interfere with the framing, a technique called **character stuffing** is used.

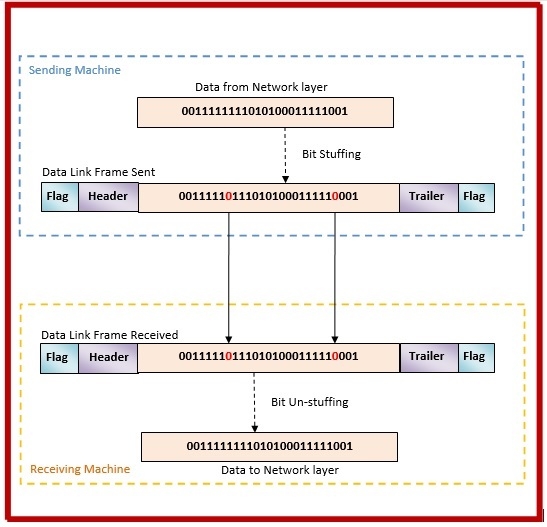
**Character stuffing** is the process where the sender’s data link layer inserts an ASCII DLE character just before the DLE character in the data. The receiver’s data link layer removes this ASCII DLE character before its given to the network layer.

However, character stuffing is closely associated with 8bit data characters and this is a major hurdle in transmitting arbitrary sized characters.

1. Starting and ending flags with **Bit stuffing** – This method allows data frames to contain arbitrary number of bits and allows character codes with an arbitrary number of bits per character.

At the start and end of the outgoing data stream, a flag byte consisting of a special bit pattern is added. Generally, the flag sequence consists of 6 or more 1s. Now, it may happen that the same flag sequency (e.g. 111111) exists in the data.

In this case, to differentiate the message from the flag sequence, the data link layer of the sender stuffs a 0 whenever it encounters 5 consecutive 1s. The receiver’s data link layer destuffs the 0 whenever it encounters 5 consecutive 1s in the data stream to restore the original message.



1. Physical layer coding violations

This method is used only for networks in which encoding on physical medium includes some sort of redundancy i.e., use of more than one graphical or visual structure to simply encode or represent one variable of data.

1. Protocol

In networking, a **protocol** is a standardized set of rules for formatting and processing data. Network protocols are like a common language for computers.

The computers within a network may use vastly different software and hardware but the use of a set of protocols enables them to communicate with each other.

E.g. whatever the configuration may be, if computer A uses Internet Protocol (IP) and computer B uses the same, they will be able to communicate with each other. However, if computer B uses a different protocol, they won’t be able to communicate with each other.

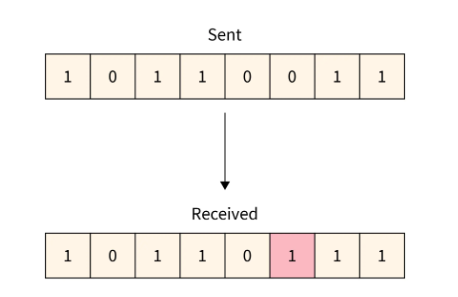
1. Error Control

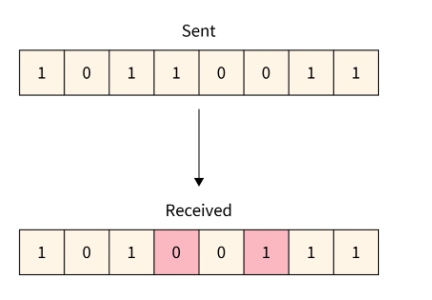
Data link layer uses techniques of error control to simply ensure and confirm that all the data frames or packets, i.e. bit streams of data, are transmitted or transferred from sender to receiver with a certain accuracy.

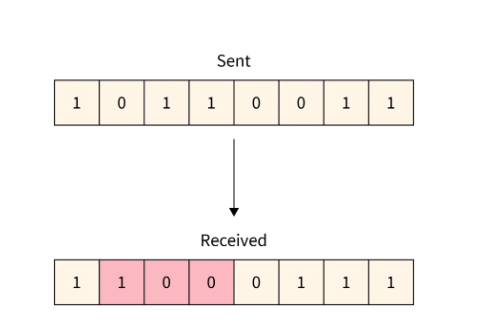
Using or providing error control at the data link layer is an optimization and never a requirement.

Types of Errors:

1. Single bit error – Typically one bit of the frame received is corrupt and the corrupted bit can be located anywhere in the frame.



1. Multiple bits error – More than one bit is corrupted and they can be located anywhere within the frame (doesn’t have to be consecutive bits).
2. 
3. Burst error - More than one bit is corrupted but they should be consecutive bits.



Block Coding:

Hamming Distance:

Hamming distance is a metric for comparing two binary data strings. While comparing two binary strings of equal length, hamming distance is the number of bit positions in which the two bits are different.

It is used for error detection or correction when data is transmitted over computer networks.

Calculation of Hamming Distance ->

For two binary strings a and b,

Hamming distance [d (a, b)] = number of 1s in (a xor b).

E.g. 11011001 ⊕ 10011101 = 01000100. Since, this contains two 1s, the Hamming distance, d(11011001, 10011101) = 2.

Error Control can be categorized into:

1. Error Detection – Identification of any discrepancy (dada) in the transmitted data.

* Parity Check
* Checksum
* Cyclic Redundancy Check (CRC)

1. Error Correction

* Hamming Code

1. Flow Control protocols

Consider a situation in which the sender transmits faster than what the receiver can accept. If the sender keeps pumping out frames at a high rate, at some point the receiver will be completely swamped and will start losing some frames. This problem may be avoided by introducing **flow control**.

Most flow control contain a feedback mechanism to inform the sender when it should transmit the next frame.

Mechanisms/Protocols for flow control:

1. Stop and Wait Protocol

It’s the simplest flow control protocol in which the sender transmits a frame and waits for acknowledgement (positive or negative) from the receiver before proceeding.

If a positive acknowledgement is received, the sender proceeds to transmit the next frame.

If a negative acknowledgement is received, the sender retransmits the same frame.

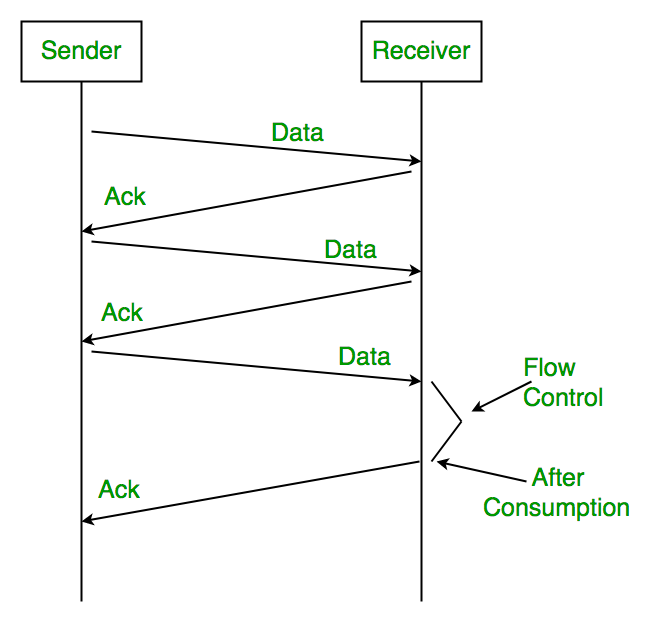
Major flaw -> If somehow a data packet or acknowledgement is completely destroyed in transit, a deadlock will occur because the sender cannot send the next frame until it receives an acknowledgement.

Problems faced by Stop and Wait ->

* Lost Data
* Lost Acknowledgement
* Delayed Acknowledgement/Data

These problems can be solved by using Stop and wait ARQ where **timers** are set up in the sender’s side. The timer starts when the frame is transmitted. If no acknowledgement is received within a certain time interval, the timer goes off and the sender may retransmit the same frame.

The ‘Stop and Wait’ protocol combined with this fix mechanism is called ‘Stop and Wait ARQ’ where ARQ is Automatic Repeat Request). It does both **flow control** and **error control**.



Advantages of Stop and Wait ARQ->

* **Simple implementation**
* **Error Detection** – It detects errors using checksums or CRC. On presence of error, the receiver sends NAK (Negative Acknowledgement) signals.
* **Reliable** – It ensures that data transmitted is reliable and are sent in order.
* **Flow Control**
* **Backward Compatibility** – Its compatible with many existing systems and protocols making it a popular choice.

Disadvantages of Stop and Wait ARQ ->

* **Low Efficiency** – Because sender has to WAIT
* **High Latency**
* **Limited Bandwidth utilization** – It doesn’t utilize the whole bandwidth since the sender can transmit only 1 frame at a time.
* **Limited error recovery** – If a data packet/frame is lost/corrupted, the sender has to retransmit the whole packet. This can be time consuming and cause further delays.

1. Sliding Window Protocol

In spite of using timers, the Stop and Wait protocol has a few more drawbacks.

(Limited bandwidth problem) + if the receiver is busy and doesn’t wish to receive any more packets, it may delay the acknowledgement but the timers will still go off and the sender perform unnecessary retransmission. These drawbacks are overcome by the Sliding window protocols.

In sliding window protocols, the sender’s data link layer maintains a ‘sending window’ which consists of a set of sequence numbers corresponding to the frames it is permitted to send.

Similarly, the receiver’s data link layer maintains a ‘receiving window’ corresponding to the set of frames it is permitted to accept.

Transmission delay (Tt) -> Time to transmit the packet from the host to the outgoing link.

**Tt = D/B [B: Bandwidth, D: Data size]**

Propagation Delay (Tp) -> Time taken by the first bit transferred by the host to the outgoing link to the destination.

**Tp = d/s [d: distance, s: wave propagation speed]**

Efficiency -> It’s the ratio of total useful time to the total cycle time of a packet.

For stop and wait protocol,

**Total Time (TT) = Tt(data) + Tp(data) + ~~Tt(ack)~~ + Tp(ack) = Tt + 2Tp** [Transmission time of ack can be neglected since acknowledgements are very less in size]

**Efficiency = Useful Time / Total Time = Tt/(Tt+2Tp) = 1/(1+2a)** [a = Tp/Tt]

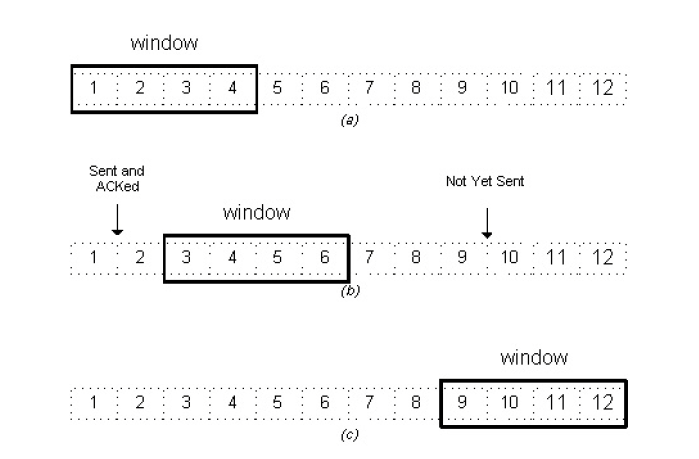
Effective Bandwidth (Throughput) -> Number of bits sent per second.

**EB = Data Size (D) / Total Time (Tt + 2Tp) = = =**

**= Bandwidth x Efficiency**

Capacity of Link -> Capacity of a link is the number of bits it can hold at maximum.

**Capacity = Bandwidth x Propagation Time = B x Tp**

****

The maximum packets that can be transmitted in total cycle time (maximum window size) = 1+2a

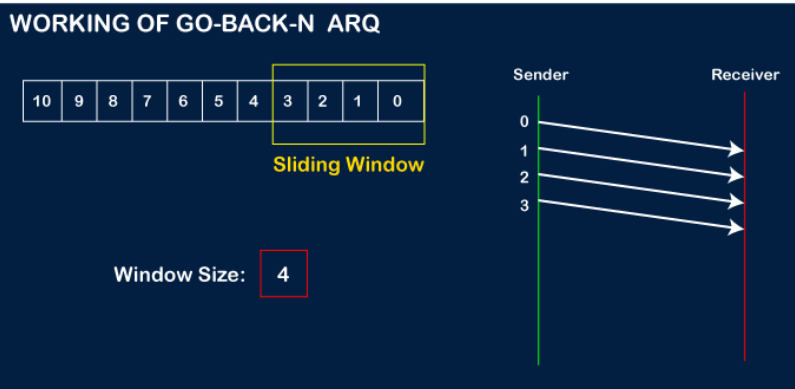
Number of bits required to represent the sender window = ceil (log2(1+2a))

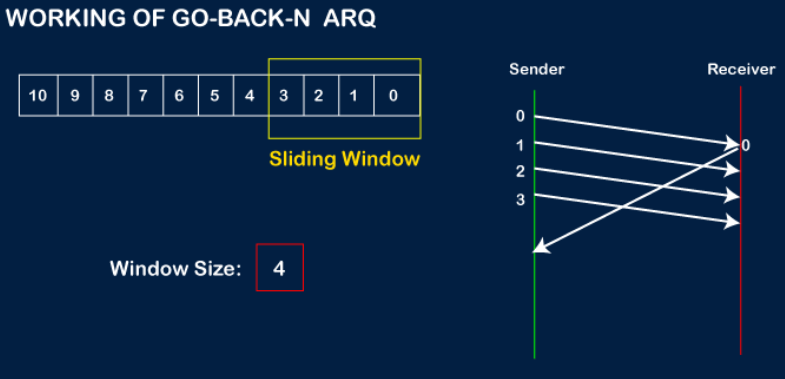
* Go Back ‘N’ ARQ

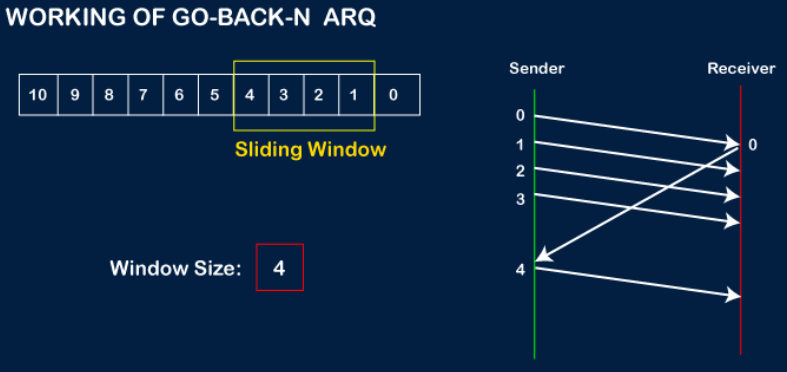
Here, the size of sender window is N. E.g. for go-back-8, the size of the sender window will be 8. **The receiver window size is always 1**.

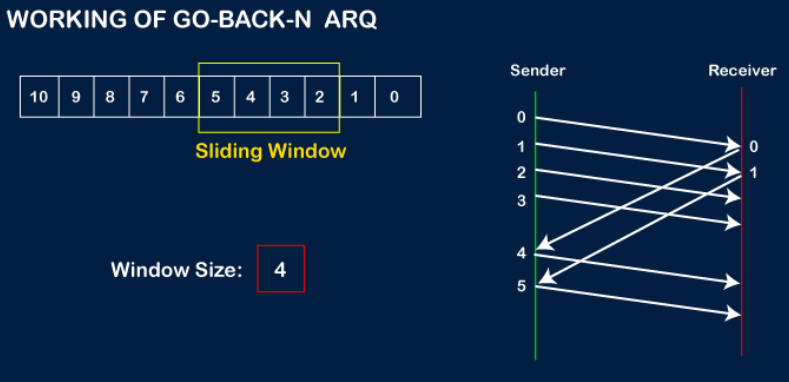
If the receiver receives a corrupted frame, it cancels it. The timer expires and the sender retransmits the same frame.

Working ->









As we can see, if a frame is lost or damaged in transmission, the receiver’s data link layer discards the subsequent frames in the window too. So, the window is shifted to the left by one frame.

* Selective Repeat ARQ

The selective repeat ARQ is a sliding window protocol strategy that is used where reliable in-order delivery of the data packets is required. The selective repeat ARQ is required for noisy channels or links and it manages the flow and error control between the sender and the receiver.

In selective repeat ARQ, we only resend the data frames that are damaged or lost. On the other hand, the correct frames are received at the receiver’s end and are buffered for future usage.

In selective repeat, the maximum window size may be calculated as ->

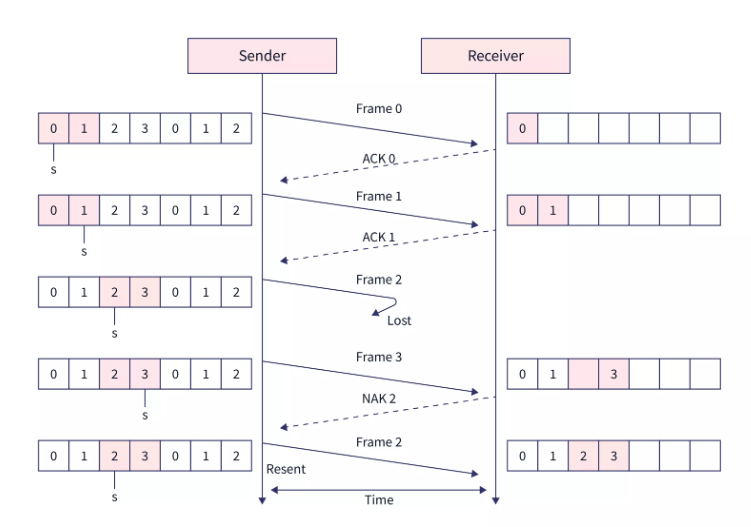
**Let sender’s window size = receiver’s window size = w**

**Initially they both contain 0 to w-1 values. Let’s assume all w frames are transmitted by sender and received by receiver correctly. However, all acknowledgements sent by receiver’s DL layer are damaged on the way.**

**Now, sender’s window is at (0, w-1) indices and receiver’s window is at (w, 2w-1) indices.**

**To avoid overlap when the sender retransmits, sum of these two windows must be less than sequence number space.**

**Max Window Size = Sequence Number Space / 2**



1. Multiple Access Control

If there is a dedicated link between the sender and receiver, the data-link control layer is sufficient. However, if there is no dedicated link present then multiple stations can access the channel simultaneously.

Hence, multiple access protocols are required to decrease collision and avoid crosstalk.

For e.g. if a teacher asks a question to a class of students (stations), all students start answering simultaneously causing chaos (data overlap/loss/collision/cross talk). It’s the teacher’s (Multiple Access Protocols) job to manage students.

1. Random Access Protocols

All stations have same superiority (No priority). Any station can send data depending on medium’s state (idle or busy).

Features of Random-Access Protocols ->

* There is no fixed time for sending data
* There is no fixed sequence of data.

Categories of Random-Access Protocols ->

* ALOHA
* CSMA
* CSMA/CD
* CSMA/CA

1. Controlled Access Protocols

* Reservation
* Polling
* Token Passing

1. Channelization Protocols

* FDMA
* TDMA
* CDMA

1. IEEE 802.3 and Ethernet

Module 4: Medium Access Sub Layer

Module 5: Network Layer

Module 3: Transport Layer

Module 4: Application Layer

Module 5: Modern Topics