Module 1

Module 2

Module 3

Module 3

Syllabus : Channel allocation problem, Multiple access Protocol (ALOHA, Carrier sense multiple access (CSMA/CD).

Module 4

Syllabus : Network Layer design issues, **Communication Primitives** : Unicast, Multicast, Broadcast.

Routing algorithms : Shortest Path (Dijkastra's), Link state routing, Distance Vector Routing.

Congestion control algorithms : Open loop congestion control, Closed loop congestion control, QoS parameters, Token and Leaky bucket algorithms.

Module 4

Syllabus : IPv4 Addressing (Classfull and Classless), Subnetting, Supernetting design problems, IPv4 Protocol, Network Address Translation (NAT), IPv6.

Protocols : ARP,RARP, ICMP, IGMP

Module 5

Syllabus : **The Transport Service**: Transport service primitives, Berkeley Sockets, Connection management (Handshake), UDP, TCP, TCP state transition, TCP timers, TCP Flow control (sliding Window), **TCP Congestion Control** : Slow Start.

□□□

Module 1

Chapter **1**

Introduction to Networking

Syllabus

Introduction to computer network, Network application, Network software and hardware components (Interconnection networking devices), Network topology, Protocol hierarchies, Design issues for layers, Connection oriented and connectionless services, Reference models : Layer details of OSI, TCP/IP models, Communication between the layers.

Chapter Contents

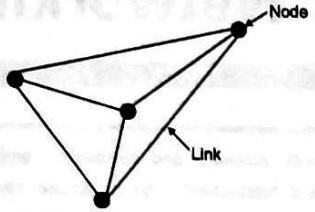
1.1 Introduction :

Network :

- Network is a broad term similar to "system". Network is a communication system which supports many users.
- The interconnection of one station to many stations is called as networking.
- A network is any interconnection of two or more stations that wish to communicate.

Node :

- Each station in a communication network is called as a node.
- The nodes are connected in different way to each other to form a network.
- One of such networks is shown in Fig. 1.1.1.



(G-13) Fig. 1.1.1 : A simple communication network

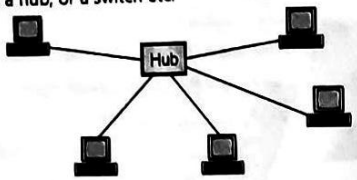
- Many other forms of interconnections are possible. The most familiar network is the telephone system.
- It is the largest and most sophisticated network of all.

1.1.1 Introduction to Computer Networks :

- A "computer network" is a system which allows communication among the computers connected in the network.
- The computers and communications have been merged together and their merger had a very deep impact on the manner in which computer systems are organized.
- In the old model a single computer used to serve all the computational needs of an organization.
- But now it is replaced by a new model in which a large number of separate but interconnected computers do the job.
- Such systems are called as **computer networks.**

Definition :

- A computer network is a group of computers and other computing hardware devices, linked together through communication channels to facilitate communication and resource sharing among wide range of users.
- The connection between the separate computers can be done via a copper wire, fiber optics, microwaves or communication satellite.
- As shown in Fig. 1.1.2, each node in a computer network is a computer, or a connecting device such as a hub, or a switch etc.



(G-1395) Fig. 1.1.2 : A computer network

- The computers connected in a network share files, folders, applications and resources like scanners, web-cams, printers etc.
- The best example of a computer network is the Internet.

1.1.2 Hardware and Software :

- In a computer network we need to make use of hardware and software.
- The **hardware** consists of connecting cables, connectors, network connecting devices and the **software** consists of protocols, programs etc.
- This enables the systematic exchange of information between the computers connected in the network.
- There are various ways of interconnecting the computers.

1.1.3 Protocol :

- For successful communication to occur, it is not enough for the "sender" to simply transmit the message and "assume" that the "receiver" will receive it properly.
- There are certain rules that must be followed to ensure proper communication.

Definition :

- Protocol is defined as the set of rules agreed upon by the sending and receiving computer systems, to facilitate a proper communication between them.
- Various protocols are used in the modern computer communication system.
- Protocols are needed to ensure proper communication among the computers connected in a computer network.

1.1.4 Components of a Computer Network :

- Following are some of the important components of a computer network :
  1. Two or more computers.
  2. Cables (coaxial, twisted pair or fiber optic) as links between the computers.
  3. A Network Interfacing Card (NIC) on each computer.
  4. Switches or other suitable connecting device.
  5. A software called network operating system.

1.1.5 Computer Network Criteria :

- Network is a broad term similar to system. Network is a communication system which supports many users.
- A network must be able to meet certain criteria. The most important of them are :

  1. Performance 2. Reliability 3. Security

1. **Performance :**
   - Performance can be measured in different ways. We can measure it in terms of transit time and response time.
   - **Transit time** is defined as the time required for a message to travel from one device to the other.
   - **Response time :** It is the time elapsed between the instant of enquiry and the instant of giving response.
   - The other factors deciding the performance are as follows :
     1. Number of users.
     2. Type of transmission medium.
     3. The hardware used.
     4. The software used.
2. **Reliability :**
   - The network reliability is important because it decides the frequency at which network failure takes place.
   - It also decides the time taken by the network to recover and its robustness in the catastrophe.
3. **Security :**
   - The network security refers to protection of data from the unauthorized user or access.
   - It also includes the data protection against damage and recovering it in the events of data losses.

1.1.6 Applications of Computer Networks :

- The computer networks are used for the following applications.
  1. For sharing the resources such as printers among all the users.
  2. For sharing of expensive softwares and database.
  3. To facilitate communication from one computer to the other.
  4. To have exchange of data and information amongst the users, via the network.
  5. For sharing of information over the geographically wide areas.
  6. For connecting the computers between various buildings of an organization.
  7. For educational purposes.

1.2 Benefits of Computer Networks :

- A network is supposed to provide its uses some unique capabilities, better than what the individual machines and their software can provide.
- The benefits provided by the network to the users can be divided into two categories as follows :
  1. Sharing
  2. Connectivity

1.2.1 Sharing Information :

- Networking allows the users to access the data stored on other's computers.
- It is possible for every user to share his bit of information with the other users over the network.
- The information sharing can be in the form of exchange of data, chatting, sending E-mails, sharing video information, groups etc.
- It is also possible for the users to share the information about various products, movies, technical information, cooking, travel books on internet.
- Sharing of information via Internet has become very common now a days.
- The information which is to be shared or being shared should be shared centrally, it must be kept consistent and secured.
- The access to this stored information should be allowed only to the authorised users.
- Sharing of information eliminates the need of transferring files on CDs or pen drives etc.

1.2.2 Sharing Resources :

- Networks can allow its users to share various types of resources.
- We can broadly categorise the shared resources as follows :
  1. Shared hardware resources
  2. Shared software resources

1. Sharing of hardware resources :

- A network allows its users to share the many hardware devices such as printers, modems, fax machines, CD ROM players etc.
- These resources are available to any one on the network irrespective of the physical location of the resource and the user.
- This will save the expenses on duplication of such hardware resources.

Sharing of software resources :

With every computer, we need to install some basic software's on each computer's hard disk.

- So each computer on the network will have to purchase a separate copy of each software required to be used. This will increase the cost to be incurred.
- In addition, installing software on each computer is time consuming and difficult.
- This problem can be overcome by using the concept of software resource sharing.
- In a network, we can centrally install and configure only one copy of each software and share it among rest of the computers. This actually saves a lot of time and cost.

1.2.3 Facilitating Centralized Management :

- The computer network facilitates centralized network management with respect to following :
  1. Management of software
  2. Maintenance of network
  3. Keeping the data back up
  4. Central network security
- All this is allowed by the **client – server** network.

1. Managing software :

- As discussed earlier, it is a very good idea to share the software resources, instead of installing a separate copy of software on each computer.
- It is possible to load all the important software on a single computer (server).
- All the other computers can make use of this centralized software as per their requirements.
- This reduces the expenses in buying the expensive software's for each individual computer. It also makes the virus checks easy.
- We can add new computers on the existing network without purchasing the software's again.
- Thus the network helps in maintaining a centralized software bank.

2. Maintenance of network :

- The second aspect in the centralized management is the maintenance of network.
- The centralized management allows quick and easy way to the routine maintenance of network.
- The client server networks are maintained centrally. It is an important but difficult job.

- A central administrator keeps track of the status of the network in respect of its speed, traffic, performance and security.
- Some of the network maintenance tools available to help the network maintenance are as follows :
  1. Protocol analyzer.
  2. Event viewer.
  3. Performance monitor.
  4. Network analyzer.
  5. Network management protocol.

3. Backing up data :

- In the process of data backup, data from computer system is copied from the disk to some other medium for keeping it safe.
- Taking back up periodically is important because it protects the data against any unpredictable, accidental loss of data due to system failure, computer viruses, or human error.
- But taking a backup of individual user's data separately is a time consuming and unorganized.
- Hence in a network, the users first save their important data on the central server and then the backup can be taken on the server data.
- This reduces the time and stores the backup data at a signal place only. This makes the data retrieval easy.
- We can have two or three sets of the entire backup data. This helps in the event of one or two sets getting corrupt.
- The duplication of backup data becomes easily possible due to centralized storage.
- The centralized backup procedures have become easy now a days due to the advanced technology.
- There are two basic network backup strategies :
  1. Isolated backup 2. Centralized backup
- The operating systems will provide tools required for data backups.
- For example windows NT proves a tape backup program called as **backup.**
- A proper backup policy which is suitable for the given network should be selected.
- Some of the backup policies are as follows :
  1. Full backup
  2. Replication
  3. Incremental or partial backup.

1.2.4 Other Benefits of Computer Networks :

- Following are some of the other advantages of computer networks.

1. Increased speed :

- Networks provide a very fast means for sharing and transfer of files.
- If the computer networks would not have been there, then we would have to copy the files on CDs or pen drive and send them to the other computers.

2. Reduced cost :

- Many popular versions of softwares usable for the entire network are now available at a considerably reduced costs as compared to individual licensed copies.
- In addition to this it is also possible to share a program on a network. It is also possible to upgrade the program.

3. Improved security :

- It is possible to protect the programs and files from illegal copying.
- By allotting password the access can be restricted to authorised users only.

4. Centralized software managements :

- Due to the use of computer networks, all the softwares can be loaded on one computer.
- All the other computers can make use of this centralized software.
- It is not necessary to waste time and energy in installing updates and tracking files on independent computers.

5. Electronic-mail :

- The computer network makes the hardware available which is necessary to install an e-mail system.
- The person to person communication is improved due to a presence of e-mail system.

6. Flexible access :

- It is possible for the authorized users to access their files from any computer connected on the network.
- This provides tremendous flexibility in accessing.

1.3 Network Services/ Applications :

- The computer networks are playing an important role in providing services to large organisations as well as to the individual common man.

1.3.1 Service Provided to Organizations :

- Many organisations have a large number of computers in operation.
- These computers may be within the same building, campus, city or different cities.
- Even though the computers are located in different locations, the organisations want to keep track of inventories, monitor productivity, do the ordering and billing etc.
- The computer networks are useful to the organisations in the following ways :

1. Resource sharing :

- It allows all programs, equipments and data available to anyone on the network irrespective of the physical location of the resource and the user.

2. High reliability due to alternative sources of data :

- It provides high reliability by having alternative sources of data.
- For e.g. all files could be replicated on more than one machines, so if one of them is unavailable due to hardware failure or any other reason, the other copies can be used.
- The aspect of high reliability is very important for military, banking, air traffic control, nuclear reactor safety and many other applications where continuous operations is a must even if there are hardware or software failures.

3. Cost :

- Computer networking is an important financial aspect for organisations because it saves money.
- Organisations can use separate personal computer one per user instead of using mainframe computer which are expensive.
- The organisations can use the workgroup model (peer to peer) in which all the PCs are networked together and each one can have the access to the other for communicating or sharing purpose.
- The organisation, if it wants security for its operation it can go in for the domain model in which there is a server and clients.
- All the clients can communicate and access data through the server.

4. Communication medium :

- A computer network provides a powerful communication medium among widely separated employees.
- Using network it is easy for two or more employees, who are separated by geographical locations to work on a report, document or R and D simultaneously i.e. on – line.

1.3.2 Services Provided to People :

- The computer networks offer the following services to an individual person :
  1. Access to remote information
  2. Person to person communication
  3. E-commerce
  4. Interactive entertainment.

1. Access to remote information :

- Access to remote information involves interaction between a person and a remote database.
- Access to remote information comes in many forms like :
- Home shopping, paying telephone, electricity bills, e-banking, on line share market etc.
- Newspaper is on-line and is personalised, digital library consisting of books, magazines, scientific journals etc.
- World wide web which contains information about the arts, business, cooking, government, health, history, hobbies, recreation, science, sports etc.

2. Person to person communication :

Person to person communication includes :

1. Electronic-mail (e-mail).

2. Real time e-mail i.e. video conferencing allows remote users to communicate with no delay by seeing and hearing each other.
3. Video-conferencing is being used for remote school, getting medical opinion from distant specialists etc.
4. Worldwide new groups in which one person posts a message and all other subscribers to the newsgroup can read it or give their feedbacks.

3. Interactive entertainment :

- Interactive entertainment includes :
  1. Multi-person real-time simulation games.
  2. Video on demand.
  3. Participation in live TV programmes likes quiz, contest, discussions etc.

1.4 Network Topology :

MU : Dec. 18, Dec. 19

University Questions

Q. 1 What is topology ? Explain the types of topologies with diagram, advantages and disadvantages. **(Dec. 18, 10 Marks)**

Q. 2 What is a topology ? Explain the types of topology. **(Dec. 19, 10 Marks)**

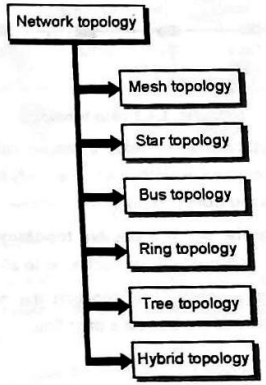
- It is possible to connect computers in many different ways in a network.
- The way of connecting the computers is called as the topology.
- So, depending on the manner of connecting the computers we can have different network topologies.

Definition :

- Topology is defined as the logical arrangement of the nodes (computers).
- The word physical network topology is used to explain the manner in which a network is physically connected.
- Devices or nodes in a network get connected to each other via communication links and all these links are related to each other in one way or the other.
- The geometric representation of such a relationship of links and nodes is known as the topology of that network.

Types :

- The five basic network topologies are as shown in Fig. 1.4.1.



(G-14(b)) **Fig. 1.4.1 : Classification of network topology**

- These topologies can be classified into two types :
  1. Peer to peer 2. Primary – secondary
- Peer to peer is the relationship where the devices share the link equally. The examples are ring and mesh topologies.
- In Primary – secondary relationship, one device controls and the other devices have to transmit through it. For example star and tree topology.

1.4.1 Bus Topology :

MU : Dec. 18, Dec. 19

University Questions

Q. 1 What is topology ? Explain the types of topologies with diagram, advantages and disadvantages. **(Dec. 18, 10 Marks)**

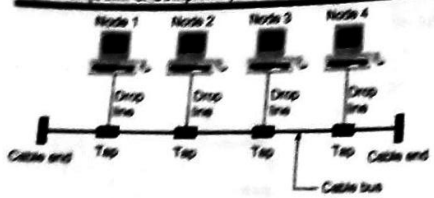
Q. 2 What is a topology ? Explain the types of topology. **(Dec. 19, 10 Marks)**

Definition :

- The bus topology is a network topology in which nodes are directly connected to the common linear or branched half duplex link called as bus.

Diagram :

- The bus topology is usually used when a network under consideration is small, simple or temporary as shown in Fig. 1.4.2.



(G-19) Fig. 1.4.2 : Bus topology

- On a typical bus network a simple cable is used without additional electronics to amplify the signal or pass it along from computer to computer.
- Therefore the bus is a **passive topology**. This long cable called bus is used as backbone to all the nodes.
- The tap is connector that connects the node to the metallic core of the bus via a drop line.

Working :

- When one computer sends a signal on the cable, all the computers on the network receive the information.
- However only the one with the address that matches with the destination address stored in the message accepts the information while all the others reject the message.
- The speed of the bus topology is slow because only one computer can send a message at a time.
- A computer must wait until the bus is free before it can transmit.
- The bus topology requires a proper termination at both the ends of the cable in order to avoid reflections.

Since the bus is a passive topology, the electrical signal from a transmitting computer is free to travel over the entire length of the cable.

Without termination when the signal reaches the end of the cable, it returns back and travels back on the cable.

The transmitted waves and reflected waves, if they are in phase add and if they are out of phase cancel. Thus addition and cancellation of wave results in a standing wave.

e standing waves can distort the normal signals ich are travelling along the cable.

- This can be avoided by terminating the bus on both ends in 50 Ω load impedance.
- The terminators absorb the electrical energy and avoid reflections.

Characteristics of the bus topology :

- Following are some of the important characteristics of the bus topology :

1. This is a multipoint configuration. There are more than two devices connected to the medium and they are capable of transmitting on the medium. Hence the Medium Access Control (MAC) is essential for the bus topology.
2. The signal strength of the transmitted signal should be adequately high so as to meet the minimum signal strength requirements of the receiver.
3. Adequate Signal to Noise Ratio (SNR) should be maintained for better quality reception.
4. The signal should not be too strong. This is necessary to avoid the overloading of transmitter and hence the possibility of signal distortion.
5. This is called as signal balancing which is not an easy task at all. Specially the signal balancing becomes increasingly difficult with increase in the number of stations.

Transmission media for bus LANs :

- We can use the following transmission media for the bus LANs :

1. Twisted pair
2. Baseband co-axial cable
3. Broadband co-axial cable
4. Optical fibre

Advantages of bus topology :

1. The bus topology is easy to understand, install, and use for small networks.
2. The cabling cost is less as the bus topology requires a small length of cable to connect the computers.
3. The bus topology is easy to expand by joining two cables with a BNC barrel connector.

4. In the expansion of a bus topology repeaters can be used to boost the signal and increase the distance.

Disadvantages of bus topology :

1. Heavy network traffic slows down the bus speed. In bus topology only one computer can transmit and other have to wait till their turn comes and there is no co-ordination between computers for reservation of transmitting time slot.
2. The BNC connectors used for expansion of the bus attenuates the signal considerably.
3. A cable break or loose BNC connector will cause reflections and bring down the whole network causing all network activity to stop.

Application :

1. Ethernet 10 base 2 also known as thinnet is an inexpensive network based on the bus topology.
2. A bus network behaves erratically if it is not terminated or improperly terminated.

1.4.2 Ring Topology :

MU : Dec. 18, Dec. 19

University Questions

Q. 1 What is topology ? Explain the types of topologies with diagram, advantages and disadvantages. **(Dec. 18, 10 Marks)**

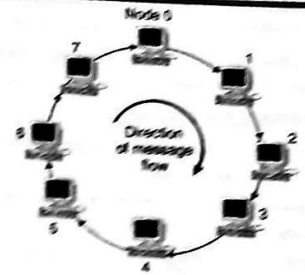
Q. 2 What is a topology ? Explain the types of topology. **(Dec. 19, 10 Marks)**

Definition :

- A ring topology is a network topology in which each node connects exactly to two other nodes, to form a single closed pathway for signal through each node.
- Data travels from node to node with each node having an access to every packet.

Diagram :

- In a ring topology, each computer is connected to the next computer, with the last one connected to the first as shown in Fig. 1.4.3.
- Rings are used in high-performance networks where large bandwidth is necessary e.g. time sensitive features such as video and audio.



(G-14) Fig. 1.4.3 : Ring topology

- Every computer is connected to the next computer in the ring and each retransmits what it receives from the previous computer hence the ring is an active network.
- The messages flow around the ring in one direction. There is no termination because there is no end to the ring.

Token passing :

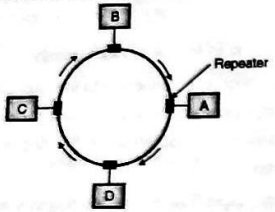
- Some ring networks do token passing. A short message called a token is passed around the ring until a computer wishes to send information to another computer.
- That computer modifies the token, adds an electronic address and data and sends it around the ring.
- Each computer in sequence receives the token and the information and passes them to the next computer until either the electronic address matches the address of a computer or the token returns to its origin.
- The receiving computer returns a message to the originator indicating that the message has been received.
- The sending computer then creates another token and places it on the network, allowing another station to capture the token and begin transmitting.
- The token circulates until a station is ready to send and capture the token.
- Faster networks circulate several tokens at once.
- Some ring networks have two counter-rotating rings that help them recover from network faults.

Active or passive ?

- Ring topology is an active topology, because each station has to recreate the packet.

Characteristics of ring topology :

1. The basic ring topology is shown in Fig. 1.4.4, which shows that along with the nodes A, B, C, D equal number of repeaters are used and that the transmission is unidirectional.



(G-17) **Fig. 1.4.4 : Ring topology**

2. The data is travels in a sequential manner around the ring.
3. Each repeater will receive regenerate and retransmit this data bit.

Problems faced in the ring topology :

1. If any link breaks or if any repeater fails then the entire network will be disabled.
2. To install a new repeater for supporting a new device, it is necessary to have the identification of two nearby, topologically adjacent repeaters.
3. It is necessary to take preventive measures to deal with the time jitter.
4. Due to the closed nature of the ring topology it is necessary to remove the circulating packets.

- These problems except for the last one can be rectified by refinements of the ring topology.

Advantages of ring topology :

1. Every computer gets an equal access to the token.
2. There are no standing waves produced.

Disadvantages of ring topology :

1. Failure of one computer on the ring can affect the whole network.
2. It is difficult to trouble short the ring.
3. Adding or removing the computers disturbs the network activity.

Application :

1. Token ring networks are defined by the IEEE 802.5 standard.
2. Fibre Distributed Data Interface (FDDI) is a fast fibre-optic network based on the ring topology.

1.4.3 Star Topology : MU : Dec. 18, Dec. 19

University Questions

Q. 1 What is topology ? Explain the types of topologies with diagram, advantages and disadvantages. **(Dec. 18, 10 Marks)**

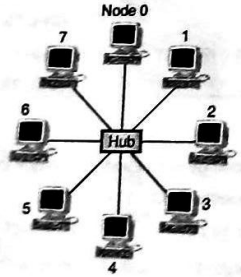
Q. 2 What is a topology ? Explain the types of topology. **(Dec. 19, 10 Marks)**

Definition :

- Star topology is a network topology, in which each individual piece of a network is connected to a central node called as a hub or switch.

Diagram :

- In a star topology all the computers are connected via cables to a central location where they are all connected by a device called a hub as shown in Fig. 1.4.5.



(G-18) **Fig. 1.4.5 : Star topology**

- There is no direct connections among the computers.
- All the connections are made via the central hub.
- Stars are used in concentrated networks, where the endpoints are directly reachable from a central location; when network expansion is expected and when the greater reliability of a star topology is needed.

- Each computer on a star network communicates with a central hub.
- The hub then resends the message to all the computers in a broadcast star network.
- It will resend the message only to the destination computer in a switched star network.
- The hub in a broadcast star network can be active or passive.
- An active hub generates the electrical signal and sends it to all the computers connected to it.
- This type of hub is usually called a multiport repeater.
- Active hubs require external power supply.
- A passive hub is a wiring panel or punch down block which acts as a connection point. It does not amplify or regenerate the signal.
- Passive hubs do not require electrical power supply.
- Several types of cables can be used to implement a star network.
- A hybrid hub can use different types of cable in the same star network.

Active or passive topology ?

- Star topology networks can be either active or passive depending on the following factors.
- If the central node performs processes like amplification or regeneration then it is an **active** topology. Otherwise it is a **passive** topology.
- If the network actively controls the data transit, then it is **active** otherwise **passive.**
- If the network requires electrical power sources then it is **active** otherwise **passive.**

Advantages :

1. It allows more devices to be attached to a single hub and can therefore increase the distance of a signal can travel between devices.
2. It allows the network to isolate and attach priorities to the communications from different computers.

Disadvantages :

1. If the central hub fails the system breaks down.
2. The cabling cost is more.

When is star topology suitable ?

- The star topology is preferred under the following circumstances :
  1. If the centralized network control is expected.
  2. If high reliability is more important than cost.
  3. If the network is to be expanded frequently.

Devices used for star topology :

- The devices used for establishing a star topology network are : twisted pair cable, or optical fiber cable, a hub or switch, suitable connectors etc.

Applications :

1. Ethernet 10 base T is a popular network based on the star topology.
2. Intelligent hubs with microprocessor that implement features in addition to repeating network. Signals provide for centralized monitoring and management of the network.

1.4.4 Mesh Topology : MU : Dec. 18, Dec. 19

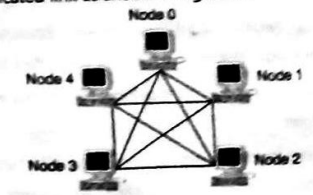
University Questions

Q. 1 What is topology ? Explain the types of topologies with diagram, advantages and disadvantages. **(Dec. 18, 10 Marks)**

Q. 2 What is a topology ? Explain the types of topology. **(Dec. 19, 10 Marks)**

Diagram :

- In a mesh topology every device is physically connected to every other device with a point to point dedicated link as shown in Fig. 1.4.6.



(G-21) **Fig. 1.4.6 : Mesh topology**

- The term dedicated means that the link carries data only between two devices connected on it.
- The mesh topology is also called as **complete topology.**

- The mesh topology does not have the traffic congestion problem, because dedicated lines are being used to connect the nodes.
- These links are not being shared.
- So the special protocol called Media Access Control (MAC) protocol is not required to be used.
- This topology has an advantage of **data security** due to the use of dedicated links. It is robust.
- If one link fails, the rest of the network can continue to function.
- The fault diagnosis and isolation of fault also is easy.
- The only disadvantages of this topology are the cable length, the cost of the cable and the associated complexity.
- A fully connected mesh network therefore has n(n–1)/2 physical cables to connect n devices.
- To accommodate that many links every device on the network must have n–1 input/output ports.
- So too many cables are required to be used for the mesh topology.
- Using this formula for a network of 1000 nodes, we will require 1000 (1000 – 1)/2 = 499500 cables or links.
- So this topology is suitable only for small networks.

Advantages :

1. The use of dedicated links guarantees that each connection can carry its own data reliably.
2. A mesh topology is robust because the failure of any one computer does not bring down the entire network.
3. It provides security and privacy because every message sent travels along a dedicated line.
4. Point to point links make fault diagnose easy.

Disadvantages :

1. Since every computer must be connected to every other computer installation and reconfiguration is difficult.
2. Cabling cost is more.
3. The hardware required to connect each link input/output and cable is expensive.

Application ;

1. Mesh topology is usually implemented as a backbone connecting the main computers of a hybrid network that can include several other topologies.

1.4.5 Tree Topology : MU : Dec. 18, Dec. 19

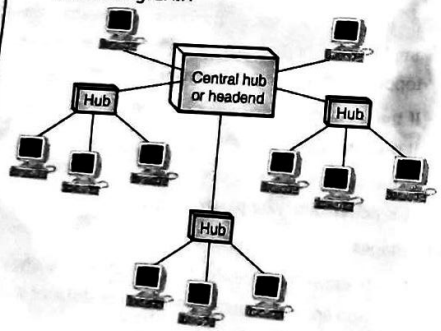
University Questions

Q. 1 What is topology ? Explain the types of topologies with diagram, advantages and disadvantages. **(Dec. 18, 10 Marks)**

Q. 2 What is a topology ? Explain the types of topology. **(Dec. 19, 10 Marks)**

Definition :

- Tree topology is a special type of structure in which many connected elements are arranged like the branches of a tree.
- A tree topology is a variation of a star.
- As in a star, nodes in a tree are connected to a central hub that controls the entire network.
- However, every computer is not plugged into the central hub.
- Most of them are connected to a secondary hub which in turn is connected to the central hub as shown in Fig. 1.4.7.



(G-22) **Fig. 1.4.7 : Tree topology**

- The central hub in the tree is an active hub which contains repeater.
- The repeater amplify the signal and increase the distance a signal can travel.
- The secondary hubs may be active or passive.

- A passive hub provides a simple physical connection between the attached devices.
- In this topology, there can be only one connection between any two nodes.
- Therefore it is also called as a parent-child topology.

Advantages :

1. It allows more devices to be attached to a single hub and can therefore increase the distance of a signal can travel between devices.
2. It allows the network to isolate and attach priorities to the communications from different computers.

Disadvantages :

1. If the central hub fails the system breaks down.
2. The cabling cost is more.

1.4.6 Logical Topology :

- Logical topology describes the manner in which the stations are logically connected to each other for the purpose of data unit exchange.
- Physical topology discussed earlier can be different from the logical topology, of the network.
- As an example consider the bus topology. The bus acts as a central controller.
- It receives data and forwards it to the various nodes.
- Thus the stations have a logical connection to the bus which acts as a centralized controller.
- Therefore the logical topology of a bus is star topology, even though the physical topology is bus.

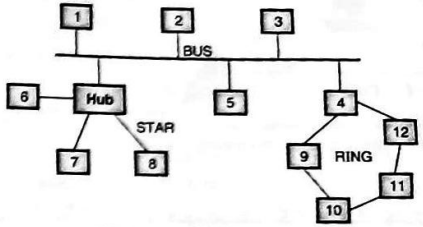
1.4.7 Hybrid Topology :

Definition :

- We have discussed various basic topologies such as bus, ring, mesh, star etc.
- Hybrid topology is the one which makes use of two or more basic topologies mentioned above, together.
- There are different ways in which a hybrid network is created.

Diagram :

- Fig. 1.4.8 shows the hybrid topology in which bus, star and ring topologies are used simultaneously.
- In Fig. 1.4.8, the nodes 1, 2, 3, 4 and 5 are connected in the bus topology, node 6, 7 and 8 form a star and the nodes 4, 9, 10, 11, 12 are arranged in a ring topology.



(G-23) **Fig. 1.4.8 : Hybrid topology**

- The practical networks generally make use of hybrid topology.
- Many complex networks can be reduced to some form of hybrid topology.
- The hybrid topology which is to be used for a particular application depends on the requirements of that application.

1.4.8 Comparison of Star, Bus and Ring Topologies : MU : Dec. 15

University Questions

Q. 1 Compare various types of network topologies. **(Dec. 15, 5 Marks)**

Sr. No.	Parameter	Bus topology	Ring topology	Star topology
1.	Configuration	Fig. 1.4.2	Fig. 1.4.4	Fig. 1.4.5
2.	Routing methodology	Only one node sends information at a time, others wait for their turn.	Information goes in one direction around the ring until it reaches the correct node.	All information passes through the control hub.
3.	Complexity	The simplest	Moderately simple	Very simple.
4.	Ease of expansion	To add a computer, we need to shut down the network.	Same as bus topology	Very easy

Sr. No.	Parameter	Bus topology	Ring topology	Star topology
5.	Reliability	Low	Low	High, depends on the reliability of central hub.
6.	Cost	Cheapest	More expensive	Most expensive
7.	Security	No security	No security	It is possible to provide security.
8.	Delay	Long	Moderate	Lowest

1.4.9 Comparison of Tree and Mesh Topologies :

Sr. No.	Parameter	Mesh topology	Tree topology
1.	Dedicated links for connections.	Used	Connections are done through hubs.
2.	Reliability	Higher	Low.
3.	Fault diagnosis	Easy	Not so easy.
4.	Number of links	Large	Small
5.	Congestion	Does not take place	Can take place.
6.	Suitable for	Small networks	Large networks.

1.5 Types of Communication :

- Based on whether the given communication system communicates only in one direction only or in both the directions, the communication systems are classified as :
  1. Simplex systems.
  2. Half duplex systems.
  3. Full duplex systems.

Simplex Communication :

- A simplex communication is defined as the communication in only one direction.
- In computer system the simplex communication takes place between CPU and keyboard or CPU and display.

Half Duplex Communication:

- A half duplex communication is defined as the bidirectional communication which does not take place simultaneously.
- These systems are bi-directional, i.e. they can transmit as well as receive but not simultaneously.
- At a time these systems can either transmit or receive, for example a transreceiver or walky talky set.
- In half duplex transmission, the entire capacity of the channel is utilized by the transmitting (sending) system.

Full Duplex Communication :

- A duplex communication is defined as the type of communication in which a simultaneous flow of information takes place at any given time.
- These are truly bi-directional systems as they allow the communication to take place in both the directions simultaneously.
- These systems can transmit as well as receive simultaneously, for example the telephone systems.
- In full duplex mode, signals going in either direction share the full capacity of link.
- The link may contain two physically separate transmission paths one for sending and another for receiving.
- Otherwise the capacity of channel is divided between signals travelling in both directions.

1.6 Network Hardware :

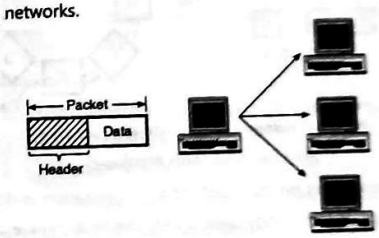
- Now let us discuss the technical issues involved in the network design.
- Two important dimensions of a computer network are :
  1. Transmission technology and 2. Scale.

1.6.1 Transmission Technology :

- The transmission technology can be categorised broadly into two types :
  1. Broadcast networks and
  2. Point-to-point networks.

1. Broadcast networks :

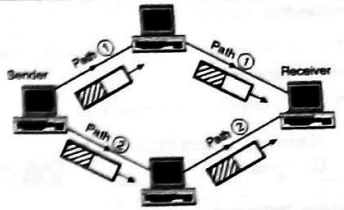
- In a broadcast networks all the machines on the network use or share communication channel that is shared or used by all the machines on the network.
- Short messages called **packets** sent by any machine are received by all the others(Fig. 1.6.1(a)).
- Broadcast systems generally use a special code in the address field for addressing a packet to all the concerned computers.
- This mode of operation is called broadcasting.
- Some broadcast systems also support transmission to only a group of few machines known as multicasting.
- When a packet is received, a machine checks the address field.
- If the packet is addressed to it then the packet is processed, otherwise the packet is ignored.
- Broadcasting is used primarily in the local area networks.



(G-2394) **Fig. 1.6.1(a) : Broadcast network**

2. Point-to-point networks :

- In point to point networks there exist of many connections between individual pairs of machines.
- To go from the source to the destination a packet on this types of network may have to go through intermediate computers before they reach the desired computer.
- As a general rule small networks which are localized in a geographical sense tend to use broadcasting (e.g. LAN) whereas networks located in wide geographical areas use point to point transmission (e.g. WAN).
- In point to point networks (Fig. 1.6.1(b)) packets can take multiple roots to reach the same destination.



(G-2395) **Fig. 1.6.1(b) : A point to point network**

- Hence routing algorithms play an important role in the point to point networks.
- Small, localized networks (e.g. LAN) tend to use the broadcasting, whereas networks located over wide geographical areas (such as WAN) use point-to-point transmission.

1.6.2 Network Scale :

- This is an alternative criterion for classification of networks.
- Fig. 1.6.1 gives the network classification based on their physical size.
- All these systems are multiprocessor systems.

Interprocessor distance	Processors are located in	Example of network
0.1 m	Same circuit board	Data flow machine
1 m	Same system	Multicomputer
10 m	Same room	LAN
100 m	Same building	LAN
1 km	Same campus	LAN
10 km	Same city	MAN
100 km	Same state	WAN
1,000 km	Same continent	WAN
10, 000 km	Same planet	Internet

Fig. 1.6.1 : Network classification according to scale

- Beyond the multi-computers are the true networks, in which the computers communicate by exchanging messages over long cables.
- Such networks are divided into following categories :
  1. Local area networks
  2. Metropolitan networks and
  3. Wide area networks.

Internetwork :

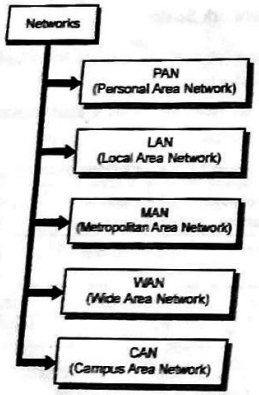
- The connection of two or more networks is called as an internetwork.
- The best example of internetwork is the Internet.

1.7 Network Classification by their Geography : MU : Dec 13

University Questions

Q. 1 What are the different categories of the network classification ? (Dec. 13, 5 Marks)

- Computer network can be classified based on the geographical area they cover, i.e. the area over which the network is spread.
- Such a classification is shown in Fig. 1.7.1.



(G-1400)Fig. 1.7.1 : Network categories

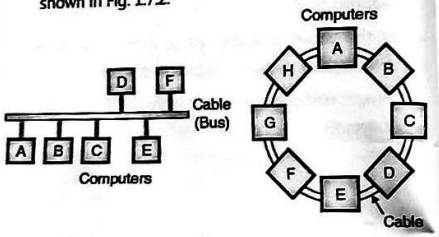
1.7.1 Local Area Networks (LAN) : MU : May 05

University Questions

Q. 1 Design a LAN network for your institute. (May 05, 6 Marks)

Definition :

- *The Local Area Network (LAN) is a network which is designed to operate over a small physical area such as an office, factory or a group of buildings.*
- *LANs are very widely used in a variety of applications.*
- *LANs are easy to design and troubleshoot.*
- The personal computers and workstations in the offices are interconnected via LAN.
- The exchange of information and sharing of resources becomes easy because of LAN.
- In LAN all the machines are connected to a single cable.
- Different types of topologies such as Bus, Ring, Star, Tree etc. are used for LANs.
- LAN uses a layered architecture and they are capable of operating at hundreds of Mbits/sec.
- A Local Area Network (LAN) is usually a privately owned and links the devices in a single office, building or campus of up to a few kilometers as shown in Fig. 1.7.2.



(a) Bus topology ***(b) Ring topology***

(G-32)Fig. 1.7.2 : LAN topologies

- *Depending on the needs of an organisation and the type of technology used, a LAN can be as simple as a few computers and a printer at home or it can contain many computers in a company and include voice, sound and video peripherals.*
- *LANs are widely used to allow resources to be shared between personal computers or workstations.*
- *The resources to be shared can be hardware like a printer or software or data.*
- *In a LAN one of the computer can become a server serving all the remaining computers called clients.*
- *Software can be stored on the server and it can be used by the remaining clients.*
- *LAN's are also distinguished from MAN's and WAN's based on the transmission media they use and topology.*

Transmission media :

- In general a given LAN will use only one type of transmission medium such as a coaxial cable or shielded twisted pair (STP).

LAN topologies :

- Different network topologies are : Bus, Ring, Star etc.
- Various topologies are possible for the broadcast LANs such as Bus topology or Ring topology.

Data rates :

- The data rates for LAN can now range from 10 Mbps to 10 Gbps.

Important characteristics / Features of LAN :

1. Very high degree of interconnection between the computers.
2. High rate of data transmission up to 10 Gbps.
3. Physical connection of computers in a LAN is easy.
4. Every computer in the LAN can communicate with every other computer.
5. The medium used for data transmission is inexpensive.

Advantages of LAN :

1. High reliability. Failure of individual computers does not affect the entire LAN.
2. It is possible to add a new computer easily.
3. The transmission of data is at a very high rate.
4. Sharing of peripheral devices such as printer is possible.

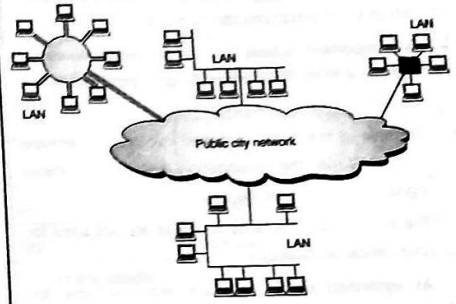
Applications of LAN :

1. File transfer and file access.
2. Personal computing.
3. Office automation.
4. Distributed computing.
5. Word and text processing.
6. Document distribution.
7. Remote access to database.
8. Electronic message handling.

1.7.2 Metropolitan Area Network (MAN) :

Definition :

- A MAN is basically a bigger version of a LAN and normally uses similar technology.
- It is designed to extend over a larger area such as an entire city.
- The MAN can be in the form of a single network such as a cable network or it can be a combination of multiple LANs as shown in Fig. 1.7.3.



(G-33)Fig. 1.7.3 : Metropolitan area network

- A MAN may be wholly owned and operated by a private company or it may be a service provided by a public company, such as a local telephone company (telco).

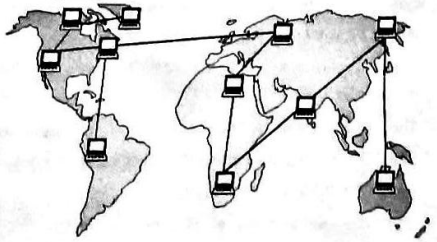
1.7.3 Wide Area Network (WAN) :

- When a network spans a large distance or when the computers to be connected to each other are at widely separated locations a local area network cannot be used.
- For such situations a Wide Area Network (WAN) must be installed.
- The communication between different users of "WAN" is established using leased telephone lines or satellite links and similar channels.
- It is cheaper and more efficient to use the phone network for the links.

Definition :

- A WAN is a telecommunications network or computer network that extends over a large geographical distance / place.
- Wide area networks are generally established with leased telecommunication circuits.
- Most wide area networks are used for transferring large blocks of data between its users.

- As the data is from existing records or files, the exact time taken for this data transfer is not a critical parameter.
- An example of WAN is an airline reservation system.
- Terminals are located all over the country through which the reservations can be made.
- It is important to note here that all the terminals use the same centralized common data provided by the central reservation computer.
- Because of the large distances involved in the wide area networks, the propagation delays and variable signal travel times are major problems.
- Therefore most wide area networks are not used for time critical applications.
- As explained earlier they are more suitable for transfer of data from one user to the other which is not a time critical application.
- Wide area networks are basically packet switching networks.
- A WAN provides long distance transmission of data, voice image and video information over large geographical areas that may comprise a country, a continent or even the whole world as shown in Fig. 1.7.4.



(G-35)Fig. 1.7.4 : Wide area network

Characteristics / features of WAN :

- Following are some of the important characteristics of WAN :
  1. Remote data entry and access is possible.
  2. Communication facility is provided.
  3. Centralized information is created and used.
  4. WAN spans over a large distance.

1.7.4 Wireless Networks :

- The fastest growing segment of the computer industry is the mobile computers such as notebook computers and Personal Digital Assistant (PDAs).
- The wireless networks are becoming increasingly important because the wired connection is not possible in cars or aeroplanes.
- Wireless networks can have many applications.

1.7.5 Internetworks :

Definition :

- When two or more networks are connected together they are called as internetwork or internet.
- Individual networks are joined into internetworks by the use of internetworking devices like bridges, routers and gateways.
- It can be a collection of number of LANs which are interconnected via a WAN.

What is the difference between a subnet and WAN ?

- If the system within a closed peripheri contains only routers then it is called as a subnet.
- But if it contains routers as well as hosts then it is a WAN.

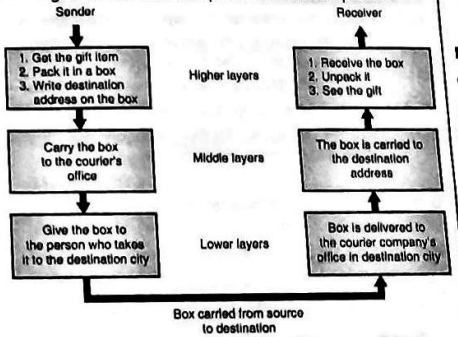
1.7.6 Comparison of LAN, WAN and MAN :

Sr. No.	Parameter	LAN	WAN	MAN
1.	Ownership of network	Private	Private or public	Private or public
2.	Geographical Area covered	Small	Very large (states or countries)	Moderate (city)

Sr. No.	Parameter	LAN	WAN	MAN
3.	Design and maintenance	Easy	Not easy	Not easy
4.	Communication medium	Coaxial cable	PSTN or satellite links	Coaxial cables, PSTN, optical fiber cables, wireless.
5.	Data rates (speed)	High	Low	Moderate
6.	Mode of communication	Each station can transmit and receive	Each station cannot transmit	Each station can transmit or receive.
7.	Principle	Operates on the principle of broadcasting	Switching	Both
8.	Propagation delay	Short	Long	Moderate
9.	Bandwidth	Low	High	Moderate

1.8 Layered Tasks :

- The concept of layers is used in our daily life. Take an example of two friends with one friend wants to send a gift to the other via courier service.
- Fig. 1.8.1 shows the steps involved in this process.



(G-1546) Fig. 1.8.1 : Layered tasks

- In Fig. 1.8.1, we have three important persons involved namely the sender, the receiver and the carrier who carries the gift box, from one city to the other.

Hierarchy of tasks :

- The point to be noted is that in order to complete a task in day to day life small actions are being done in a hierarchical way or layered manner.

1. At the sender :

Upper layers :

- The tasks of higher layers :
  1. Get the gift item
  2. Pack it in a box
  3. Write the destination address on the box.

Middle layer : Carry the addressed box to the office of a courier company.

Lower layer : Give the box to a person who will take it to the destination city.

2. At the receiver :

Tasks of lower layers :

- The box is delivered to the courier company office in the destination city.

Middle layers :

- The box is carried by another person to the destination address and the box is delivered.

Upper layers :

1. Receive the box 2. Unpack it 3. See the gift

Hierarchy and layered tasks :

- This discussion demonstrates that the important tasks are carried out by the higher layers whereas the simpler tasks are carried out by the middle and lower layers.
- In the network protocols as well the layered architecture is used.

1.9 Network Software :

- The software used in networks is equally important as the hardware.
- The network software is highly structured now a days.

1.9.1 Protocol Hierarchies (Layered Architecture) :

MU : May 12, Dec. 14, Dec. 17

University Questions

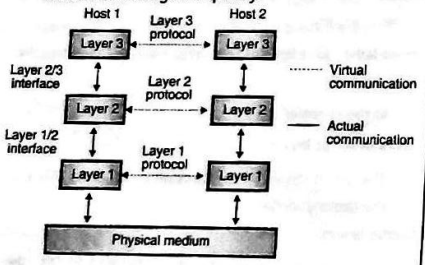
Q. 1 Draw the layered structures and compare the two network reference models – OSI and TCP/IP. **(May 12, 10 Marks)**

Q. 2 Why there is a need for layered designing for networking and communication ? Compare the TCP/IP and OSI reference models. **(Dec. 14, 10 Marks)**

Q. 3 Explain the need of layered design for communication and networking. Compare the OSI reference model & TCP/IP. **(Dec. 17, 10 Marks)**

Block diagram :

- Most networks are organized in the form of a series of layers or levels as shown in Fig. 1.9.1 in order to reduces the design complexity.



(G-49) **Fig. 1.9.1 : Layers, protocols and interfaces**

- The number of layers, the name of each layer, the contents of each layer and the function of each layer differ from network to network.
- The purpose of each layer is to offer certain services to the higher layers.
- Layer n on one machine (source) will communicate with layer n on another machine (destination).
- The rules and conventions used in this communication are collectively known as the layer "n" protocol.
- A protocol is an agreement between the two communicating machines about how the communication link should be established, maintained and released.
- Violation of the protocol will lead to the communication difficulties or failure.

Peer :

- A three layer network is shown in Fig. 1.9.1. The entities comprising the corresponding layers on different machines are called as **peers.**
- The communication actually takes place between the peers using the protocol.
- The dotted lines in Fig. 1.9.1 shows the virtual communication and physical communication is shown by solid lines.

1.9.2 Reasons for having Layered Protocols and its Benefits :

MU : May 05, Dec. 05, May 07, Dec. 09, May 13, Dec. 19

University Questions

Q. 1 Explain the need for the layered architecture of computer network. Also explain data transmission and protocols in layered architectures. **(May 05, Dec. 05, May 07, 10 Marks)**

Q. 2 Explain the need for the layered architecture in computer network. Explain how information is exchanged between two nodes using OSI model. **(Dec. 09, 10 Marks)**

Q. 3 What is the need for layering ? Discuss the design issues for layers. **(May 13, 10 Marks)**

Q. 4 Explain the need of layering for communication and networking. **(Dec. 19, 5 Marks)**

- The process of establishing a link between two devices to communicate and share information is complicated.
- There are many functions which are to be taken into consideration to allow an effective communication to take place.
- To organize all these functions in an organized way the designers felt the need to develop network architecture.
- In the network architecture various tasks and functions are grouped into related and manageable sets called **LAYERS.**

- A network architecture can be defined as a set of protocols that tell how every layer is to function.
- The reasons and advantages of using the network architecture are as follows :

1. It simplifies the design process as the functions of each layers and their interactions are well defined.
2. The layered architecture provides flexibility to modify and develop network services.
3. The number of layers, names of the layers, and the tasks assigned to them may change from network to network. But for all the networks, always the lower layer offers some services to its upper layer.
4. The concept of layered architecture is a new way of looking at the networks.
5. Addition of new services and management of network infrastructure becomes easy.
6. Due to segmentation (layered structure), it is possible to break difficult problems into smaller and more manageable tasks.
7. Logical segmentation allows parallel working by different teams on different tasks simultaneously.

1.9.3 Disadvantages of Layered Architecture :

1. The problem associated with the layered protocols is that we loose touch with the reality.
2. Layering is a kind of hiding information.
3. Layered architecture can sometimes result in poor performance.

1.9.4 How does Data Transfer take Place ?

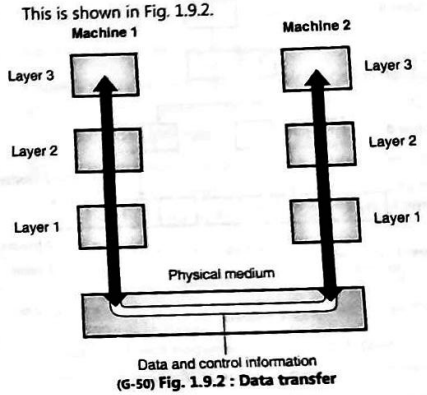
MU : May 10, Dec. 12, May 13

University Questions

Q. 1 What is the difference between protocol and service interface ? Explain your answer in terms of the OSI seven layer model. **(May 10, 10 Marks)**

Q. 2 Differentiate between protocol and interface ? **(Dec. 12, May 13, 10 Marks)**

- Data does not get transferred directly from layer n of one machine to layer n of the other machine.
- Instead the data transfer takes place as explained below.
- The data and control information is passed on to the lower layers until the lowest layer (layer 1) is reached.
- Below layer 1 lies the physical medium such as coaxial cable, through which the actual transfer of data and control information takes place.
- This is shown in Fig. 1.9.2.



(G-50) **Fig. 1.9.2 : Data transfer**

Interface :

- An interface defines the operations and services offered by lower layer to the upper layer.
- There is an interface between each pair of adjacent layers.

1.10 Network Architecture :

Definition :

- A set of layers and protocols is called as network architecture.
- Protocol stack is defined as a list of protocols used for a certain system, one protocol per layer.

1.10.1 Virtual Communication between Layers : MU : May 05, Dec. 05, May 07

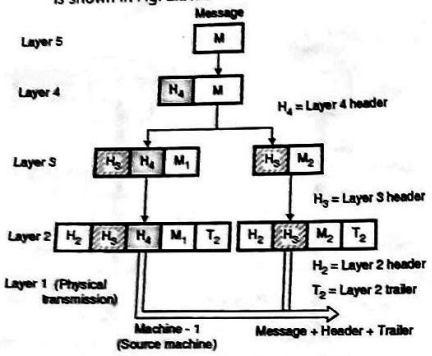
University Questions

Q. 1 Explain the need for the layered architecture of computer network. Also explain data transmission and protocols in layered architectures. **(May 05, Dec. 05, May 07, 10 Marks)**

- Let us now go into technical details of the communication between say layer 5 of two machines.
- Let machine-1 be the sending machine while machine-2 be the receiving one.

Sequence of operations at Machine-1 :

- The sequence of operation taking place at machine-1 is shown in Fig. 1.10.1.



(G-51) **Fig. 1.10.1 : Information flow for virtual communication at Machine-1**

Step 1 :

- A messages M is produced by layer -5 of machine- 1 and given to layer 4 for transmission.

Step 2 :

- Layer-4 adds a header H_4 in front of the message so as to identify the message and passes the (header + message) to layer-3.
- A header includes the control information and it allow a layer 4 in machine 2 to deliver the messages in right order.

Step 3 :

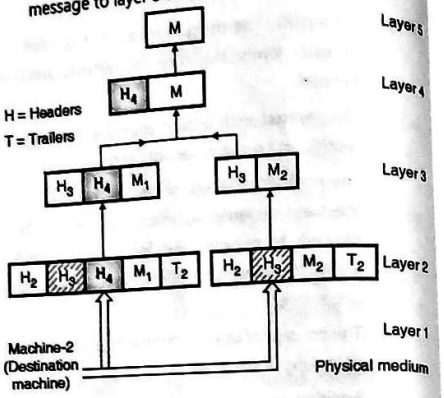
- Layer-3 breaks up the incoming messages into small units, packets and appends a layer-3 adder to each packet M_1 and M_2 as shown in Fig. 1.10.1 and passes these packets to layer-2.

Step 4 :

- Layer-2 adds header as well as trailer to each packet obtained from layer-3 and hands over the resultant unit to layer-1 for physical transmission

Sequence of operations at Machine-2 :

- The control information placed in headers is used at the destination machine (machine-2) to convey the message to layer 5 as shown in Fig. 1.10.2.



(G-52) **Fig. 1.10.2 : : Information flow for virtual communication at Machine-2**

- The sequence of events taking place at the receiving machine is exactly opposite to that taking place at the sending one.
- This process is also known as encapsulation and decapsulation.

1.11 Design Issues for the Layers :

MU : Dec. 10, May 13, May 15, Dec. 15, Dec. 16, Dec. 18, May 19

University Questions

Q. 1 Describe any five design issues for the layers. **(Dec. 10, 5 Marks)**

Q. 2 What is the need for layering ? Discuss the design issues for layers. **(May 13, 10 Marks)**

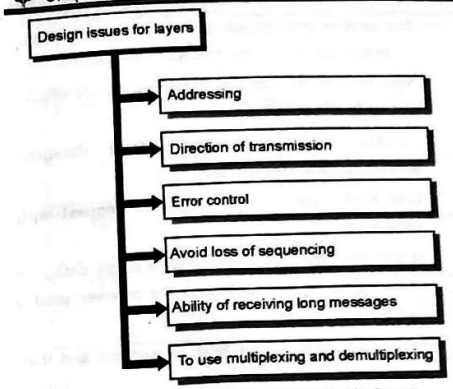
Q. 3 What are the design issues for the layers ? **(May 15, Dec. 18, 4 Marks)**

Q. 4 Discuss the design issues for various layers. **(Dec. 15, 5 Marks)**

Q. 5 List design issues in OSI layers. **(Dec. 16, 5 Marks)**

Q. 6 Explain design issues of layers. Explain ISO OSI reference model with diagram. **(May 19, 10 Marks)**

- In this section we are going to discuss some of the important design issues that are related to computer networking.
- Fig. 1.11.1 enlists various design issues for the layers.



(G-2839) **Fig. 1.11.1 : Design issues for the layers**

1. Addressing :

- For every layer, it is necessary to identify senders and receivers.
- Some mechanism needs to be used for the same.
- Since there are many possible destinations for a packet, some form of addressing is needed in order to specify a specific destination.

2. Direction of Transmission :

- Another important design issues is the direction of data transfer.
- Depending on the ability of a system to communicate only in one direction or both the directions, the communication systems are classified as :
  1. Simplex systems.
  2. Half duplex systems.
  3. Full duplex systems.

3. Error control :

- Another important issue is the error control because physical communication channels can introduce errors in the data travelling on them.
- Error detection and correction both are essential.
- Many error detecting and correcting codes are known out of which those which are agreed upon and receiver should be used.
- The receiver should be able to tell the sender by some means, that it has received a correct message or a wrong message.

4. Avoid loss of sequencing :

- All the communication channels cannot preserve the order in which messages are sent on it.
- So there is a possibility of loss of sequencing. That means messages are not received serially at the receiver.
- To avoid this, all the packets of a message should be numbered so that they can be put back together at the receiver in the appropriate sequence.

5. Ability of receiving long messages :

- At several levels, one more problem needs to be solved, which is inability of all processes to accept arbitrarily long messages.
- So a mechanism needs to be developed to deassemble (break into small messages), transmit and then reassemble messages.

6. To use multiplexing and demultiplexing :

- Multiplexing and demultiplexing is to be used to share the same channel by many sources simultaneously.
- It can be used for any layer. Multiplexing is needed at the physical layer level.

1.12 Connection Oriented and Connectionless Services :

MU : Dec. 03, May 04, Dec. 11

University Questions

Q. 1 Write advantages and disadvantages of connection oriented service with connection less service and write example application. **(Dec. 03, 10 Marks)**

Q. 2 What are the principle differences between connectionless and connection oriented communication ? Characterize all the aspects in terms of quality and service. **(May 04, Dec. 11, 10 Marks)**

- Any layer can offer two types of services to the layer above it.
  1. Connection oriented service
  2. Connectionless service.

1. Connection oriented service :

- The connection oriented service is similar to the one provided in the telephone system.

- The service users of the connection oriented service undergo the following sequence of operation :
  1. Establish a connection.
  2. Use the connection.
  3. Release the connection.
- The connection acts like a tube. The sender pushes bits from one end of the tube and the receiver takes them out from the other end.
- The order is generally preserved. That means the order in which the bits are sent is same as the order in which they are received.
- Sometimes after establishing a connection, the sender and receiver can discuss and negotiate about parameters to be used such as maximum message size, quality of service and some other issues.

2. Connectionless service :

- The connectionless service is similar to the postal service.
- Each message (analogous to a letter) carries the full address of the destination.
- Each message is routed independently from source to destination through the system.
- It is possible that the order in which the messages are sent and the order in which they are received may be different.

Quality of Service (QoS) :

- Each service can be judged by its quality of service.
- Services can be of two types :
  1. Reliable 2. Unreliable.
- Reliable services are those which never lose data. In the reliable services a receiver sends acknowledgements of the received messages to the sender.

But due to acknowledgements the overheads and delays increase which are sometimes undesirable.

Applications such as electronic mail do not require any connections.

The cost associated, complexity and overheads of reliable services is not required here.

Such applications require high reliability of message arrival but no guarantee i.e. unreliable service will be acceptable for this application.

- The services in which acknowledgements are not sent to sender are unreliable connectionless services.
- Such services are called as **datagram** service which is similar to telegram service.
- However note that **acknowledged datagram service** can also be provided.
- One more type of service is the **request-reply service.**
- In this type, the sender transmits a single datagram which contains a request and the receiver send a reply to it.
- Table 1.12.1 lists various types of services and their examples.

Table 1.12.1 : Six different types of services

Sr. No.	Service	Type	Example
1.	Reliable message stream.	Connection oriented	Sequence of pages.
2.	Reliable byte stream.	Connection oriented	Remote login.
3.	Unreliable connection.	Connection oriented	Digitized voice.
4.	Unreliable datagram.	Connectionless	Electronic mail.
5.	Acknowledged datagram.	Connectionless	Registered e-mail.
6.	Request-Reply.	Connectionless	Database query.

- The unreliable service is used only if the reliable service is not available or is too costly to afford.

1.12.1 Examples of C.O. and C.L. Services:

- Table 1.12.2 lists some of the important protocols and their categories.

Table 1.12.2

Sr. No.	Name	Type	Comments
1.	PSTN	C.O.	Uses dedicated physical connections
2.	ATM	C.O.	Provides service through virtual circuit connections

Sr. No.	Name	Type	Comments
3.	Frame Relay	C.O.	Provides service through virtual circuit connections
4.	IP	C.L.	Packet routing is used (unreliable service)
5.	TCP	C.O.	Provides a reliable service over IP protocol
6.	UDP	C.L.	Provides connectionless unreliable service over IP protocol

1.12.2 Comparison of C.O and C.L. Services :

MU : Dec. 12, May 13, May 16, Dec. 18

University Questions

Q. 1 Differentiate between the connectionless and connection oriented service. **(Dec. 12, May 13, Dec. 18, 10 Marks)**

Q. 2 Compare connection oriented and connectionless services. **(May 16, 4 Marks)**

Table 1.12.3 : Comparison of C.O and C.L. Services

Sr. No.	Parameter	Connection oriented	Connectionless
1.	Reservation of resources	Necessary	Not necessary
2.	Utilization of resources	Less	Good
3.	State information	Lot of information required	Not much information is required to be stored
4.	Guarantee of service	Guaranteed	No guarantee
5.	Connection	Connection needs to be established	Connection need not be established
6.	Delays	More	Less
7.	Overheads	Less	More
8.	Packets travel	Sequentially	Randomly
9.	Congestion due to overloading	Not possible	Very much possible

1.13 Interface and Services :

- The basic function of each layer in the layered structure is to provide service to the layer above it.
- Refer Fig. 1.13.1 which shows two layers n and (n+1).
- Thus layer n provides services to the layer (n+1).

Entities and Peer Entities :

- An entity is defined as the active element in each layer. An entity can be either a software entity or a hardware entity.
- The example of software entity is a process and that of a hardware entity is an intelligent I/O chip.
- Entities in the same layer but on different machines are called as **peer entities.**

Service Provider and Service User :

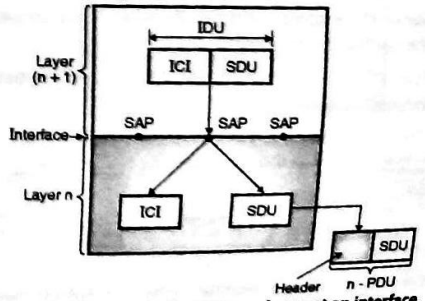
- The entities at layer n implement services for the layer (n + 1) which is above the n^{th} layer.
- So, layer n which provides service is called as service provider and layer (n + 1) which takes this service is called as service user.

Interface :

- In computing, a network interface is a software or hardware interface between two pieces of equipment or protocol layers in a computer network.
- A network interface will usually have some form of network address.
- This may consist of a node identifier and a port number or may be a unique node ID in its own right.
- Network interfaces provide standardized functions such as passing messages, connecting and disconnecting, etc.

Service Access Points (SAPs) :

- Refer Fig. 1.13.1 to understand the definition of SAPs.
- The long form of SAP is service access point. They are available at the interface of n and (n + 1) layer as shown in Fig. 1.13.1.

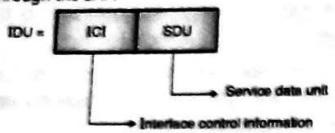


(G-56) Fig. 1.13.1 : Relation between layers at an interface

- Services are available at SAPs.
- That means the layer n SAPs are those places at the interface where layer (n + 1) can access the services being offered.
- Each SAP has a unique address for its identification.

Interface Data Unit (IDU) :

- For successful exchange of information between two layers, a set of rules about the interface should be present.
- As shown in Fig. 1.13.2, the layer (n + 1) entity passes an IDU (interface data unit) to the layer n entity through the SAP.



(G-57) **Fig. 1.13.2 : IDU**

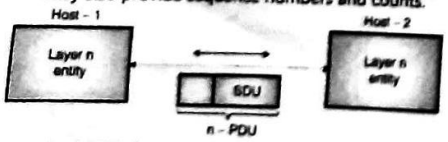
- An IDU consists of two parts namely SDU (service data unit) and ICI (interface control information).

Service Data Unit (SDU) :

- SDU is a part of IDU. The SDU is the information passed across the network to the peer entity and then up to layer (n + 1).
- ICI contains the control information which is necessary to help the lower layer (n) to do the necessary job.

Protocol Data Unit (PDU) :

- In order to transfer the SDU, the layer n entity has to divide it into many smaller pieces.
- Each piece is given a header and sent as a separate PDU (Protocol Data Unit) such as a packet.
- The PDU headers are used by the peer entities to carry out their peer protocol.
- Some PDUs contain data while other PDUs contain the control information.
- The PDU headers will identify or differentiate between different types of PDUs.
- They also provide sequence numbers and counts.



(G-58) **Fig. 1.13.3 : Layer n entities exchange n-PDUs in their layer n protocol**

1.13.1 Service :

Definition :

- It is defined as a set of operations that a layer can provide to the layer above it.
- A service defines or states the operations a layer is ready to perform.
- But it does not say anything about how these operations would be implemented.

1.13.2 Protocol :

MU : May 10, Dec. 12, May 13

University Questions

Q. 1 What is the difference between protocol and service interface ? Explain your answer in terms of the OSI 7 layer model. **(May 10, 10 Marks)**

Q. 2 Differentiate between protocol and interface ? **(Dec. 12, May 13, 10 Marks)**

- A protocol is a set of rules. The format and meaning of frames, packets or messages that are being sent and received by the communicating peer entities is governed by the protocols.
- The entities use protocols so as to implement their services.
- Once their predefined services are ensured, they are free to change the protocol.

Difference between Interface and Protocol :

- An interface refers to a point of interaction between components, whether a piece of hardware such as a graphics card or a piece of software such as an internet browser to function independently while communicating with other components via protocol.
- A protocol is a convention or standard that controls or enables the connection, communication, and data transfer between computing endpoints, or interfaces.

Difference between Service and Protocol :

- A **protocol** is a set of rules for communication within a layer.
- A **service** is what the layer provides to the layer above it through an **interface. Protocols** at one layer are unaware of issues at another layer.

1.14 Reference Models :

MU : May 17

University Questions

Q. 1 What is ISO-OSI reference model ? Compare it with TCP/IP reference model. Which layer is used for the following :

1. To route packets
2. To convert packets to frame
3. To detect and correct errors.
4. To run services like FTP, telnet etc.

(May 17, 10 Marks)

Q. 2 Describe OSI reference model with a neat diagram. **(May 11, 10 Marks)**

Q. 3 Explain the layer details of OSI and TCP/IP models. **(Dec. 12, 10 Marks)**

Q. 4 Describe the OSI Reference Model with a neat diagram. **(May 15, 10 Marks)**

Q. 5 Explain design issues of layers. Explain ISO OSI reference model with diagram. **(May 19, 10 Marks)**

Definition :

- In computer networks, reference models give a conceptual framework that standardizes communication between heterogeneous networks.

Types of reference models :

- The two most important reference models are :
  1. The OSI reference model and
  2. The TCP/IP reference model.
- The International Standards Organisation (ISO) covers all aspects of network communication in the Open Systems Interconnection (OSI) model.
- An OSI model is a layered framework for the design of network systems that allows for communication across all types of computer systems.
- The purpose of each layer is to offer certain services to the higher layers.
- Layer n on one machine (source) will communicate with layer n on another machine (destination).
- The rules and conventions used in this communication are collectively known as the layer n protocol.
- Basically a protocol is an agreement between the two communicating machines about how the communication link should be established, maintained and released.

1.15 OSI Model :

MU : May 10, May 11, Dec. 12, May 15, May 19

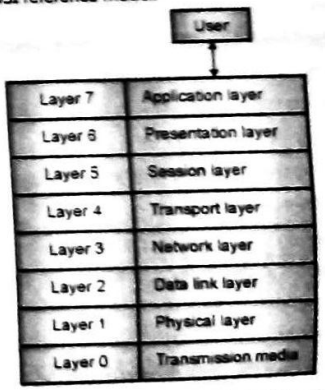
University Questions

Q. 1 What is the difference between protocol and service interface ? Explain your answer in terms of the OSI 7 layer model. **(May 10, 10 Marks)**

- The users of a computer network are located over a wide physical range i.e. all over the world.
- Therefore to ensure that nationwide and worldwide data communication systems can be developed and are compatible to each other, an international group of standards has been developed.
- These standards will fit into a framework which has been developed by the "International organization of standardization (ISO)".
- This framework is called as "Model for open system interconnection (OSI)" and it is normally referred to as "OSI reference model".

1.15.1 Layered Architecture :

- Fig. 1.15.1 shows the seven layer architecture of ISO-OSI reference model.



(G-59) **Fig. 1.15.1 : A seven layer ISO-OSI reference model**

- It defines seven levels or layers in a complete communication system.
- The lowest layer is physical layer and highest one is called as the application layer.

- Each computer on a network uses a series of protocols to perform the functions assigned to each layer.
- These layers collectively form the protocol stack or networking stack.
- At the top of the stack we have the application and at the bottom is the physical medium which actually connects the computers to form a network.

Who developed the OSI model ?

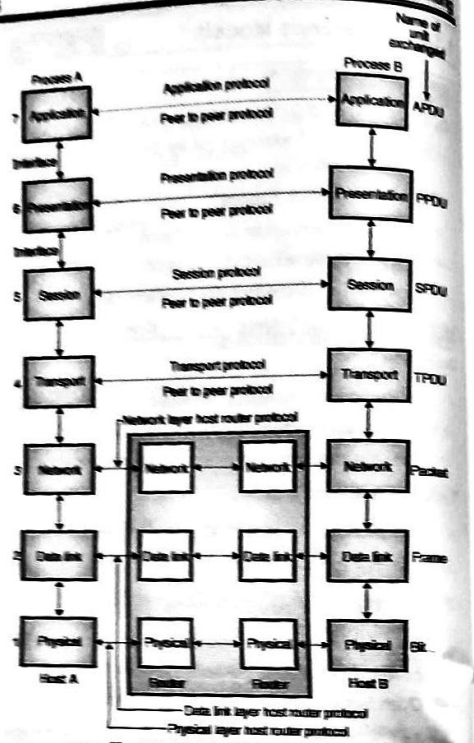
- The OSI model was developed in two different and completely independent projects by the International Organization for Standardization (ISO) and the Consultative Committee for International Telephone and Telegraphy (CCITT) which is now known as ITU-T.
- Both these organizations developed their own seven layer models.
- Then in 1983 the two projects were combined together.

OSI protocol suite :

- The OSI model was originally developed as a model for creating a 7 layer protocol suite.
- But this seven layer protocol suite never came into existence.
- In fact none of the protocol suites existing today exactly match the seven layer structure of the OSI model.
- But still the OSI reference model is so simple yet powerful that it is being used as a teaching, reference and communications tool.
- The reason why real protocol stacks differ from the OSI model is that many protocols used today were developed before the OSI model was developed.
- The TCP/IP protocols which are used extensively in practice have their own layered model.
- The TCP/IP reference model is discussed later on.

1.15.2 Communication in OSI Model :

- Fig. 1.15.2 shows a more detailed OSI model with two hosts A and B communicating with each other.



(6-68) Fig. 1.15.2 : The OSI reference model

Interface :

- An interface defines the operations and services offered by lower layer to the upper layer.
- There is an interface between each pair of adjacent layers as shown in Fig. 1.15.2.

Peer :

- The active elements present in each layer are known as **entities.**
- The entities can be hardware entities or software entities.
- The entities comprising the corresponding layers on different machines are called as **peers.**
- The communication actually takes place between the peers using the protocol.
- The dotted lines in Fig. 1.15.2 show the virtual communication and physical communication is shown by solid lines.
- Within a single machine, each layer uses the services of the layer just below it.

- However between two machines A and B of Fig. 1.15.2 layer x on machine A will communicate with layer x on machine B.
- This communication is based on some mutually agreed rules called **protocols.**
- The processes on each machine which communicate at a given layer are called as **peer-to-peer processes.**

1.15.3 Peer to Peer Processes :

- All the applications need not use all the seven layers shown in Fig. 1.15.1.
- The lower three layers are enough for most of the applications.
- Each layer is built from electronic circuits and/or software and has a separate existence from the remaining layers.
- Each layer is supposed to handle message or data from the layers which are immediately above or below it.
- This is done by following the protocol rules. Thus each layer takes data from the adjacent layer, handles it according to these rules and then passes the processed data to the next layer on the other side.

Interlayer Communication :

- In order to get an idea of interlayer communication, let us take a simple example first.
- We want the data to get transferred from layer-3 of machine-A to layer 3 of machine-B.
- But the data does not get transferred directly from layer 3 of one machine to layer 3 of the other machine.
- The data and control information is passed on from the topmost layer to the lower layers until the lowest layer (layer 1) is reached.
- Below layer 1 lies the physical medium such as coaxial cable, through which the actual communication takes place.
- This is shown in Fig. 1.15.2.
- This is called as the actual communication between the layers.

1.15.4 Organization of the Layers :

- The seven layers in the OSI model can be considered to belong to three subgroups as follows :

1. **Subgroup 1 : First three layers :**
   - The first subgroup consists of layers 1, 2 and 3 i.e the physical, data link and network layers.
   - They are important for the physical aspects of moving data from one computer to the other.
2. **Subgroup 2 : Layers 5,6 and 7.**
   - The second subgroup is made up of the upper three layers (5, 6 and 7) i.e. session, presentation and application layers.
   - This is called as the user support layers.
   - They allow the interaction between unrelated software systems.
3. **Subgroup 3 : Transport layer :**
   - The third subgroup consists of only the fourth layer i.e. the transport layer.
   - It links the subgroups 1 and 2.
   - The upper layers are implemented using software only whereas the lower layers are a combination of hardware and software.
   - The **physical layer** is implemented by only hardware.
   - Table 1.15.1 shows various layers and its functions.

Table 1.15.1 : Functions of the layers of ISO-OSI model

Level	Name of the layer	Functions
1.	Physical layer	Make and break connections, define voltages and data rates, convert data bits into electrical signal. Decide whether transmission is simplex, half duplex or full duplex.
2.	Data link layer	Synchronization, error detection and correction. To assemble outgoing messages into frames.
3.	Network layer	Routing of the signals, divide the outgoing message into packets, to act as network controller for routing data.

Level	Name of the layer	Functions
4.	Transport layer	Decides whether transmission should be parallel or single path, multiplexing, splitting or segmenting the data, to break data into smaller units for efficient handling.
5.	Session layer	To manage and synchronize conversation between two systems. It controls logging on and off, user identification, billing and session management.
6.	Presentation layer	It works as a translating layer.
7.	Application layer	Retransferring files of information, LOGIN, password checking etc.

1.15.5 Layer Details of OSI Model :

MU : May 17

University Questions

Q. 1 What is ISO-OSI reference model ? Compare it with TCP/IP reference model. Which layer is used for the following :

1. To route packets
2. To convert packets to frame
3. To detect and correct errors.
4. To run services like FTP, telnet etc.

(May 17, 10 Marks)

Layer 1 : The physical layer :

- Functions of the physical layer are as follows :
  1. To activate, maintain and deactivate the physical connection.
  2. To define voltages and data rates needed for transmission.
  3. To convert the digital data bits into electrical signal.
  4. To decide whether the transmission is simplex, half duplex or full duplex.
- A physical layer does not perform the following operations :
  1. It does not detect or correct errors.
  2. It does not decide the medium or modulation.
- The examples of the physical layer protocols are RS-232 or RS-449 standards.

Layer 2 : Data link layer :

- Functions of the data link layer are synchronization and error control for the information which is to be transmitted over the physical link.
- To enable the error detection, it adds error detection bits to the data which is to be transmitted.
- The encoded data is then passed to the physical layer.
- These error detection bits are used by the data link layer on the other side to detect and correct the errors.
- At this level the outgoing messages are assembled into frames, and the system waits for the acknowledgements to be received after every frame transmitted.
- Correct operation of the data link layer ensures reliable transmission of each message.
- Following are the functions of data link layer :
  1. Framing
  2. Flow control :
  3. Physical addressing :
  4. Error control :
  5. Access control :

Protocols at D.L.L. :

- Examples of data link layer protocols are HDLC, SDLC and X.25 protocols.

Layer 3 : The network layer :

- The functions of network layer are as follows :
  1. To route the signals through various channels to the other end.
  2. To act as the network controller by deciding which route data should take.
  3. To divide the outgoing messages into packets and to assemble incoming packets into messages for the higher levels.
  4. It translates logical network address into physical machine addresses i.e. the numbers used as destination IDs in the physical network cards.

5. It determines the quality of service by deciding the priority of message and the route a message will take if there are several ways a message can get to its destination.
6. It breaks the larger packets into smaller packets if the packet is larger than the largest data frame the data link will accept.
7. It is concerned with the circuit, message or packet switching.
8. It provides connection oriented services, including network layer flow control, network layer error control and packet sequence control.

Layer 4 : Transport layer :

- As the name suggests this layer provides the transport services. The functions of the transport layer are as listed below :
  1. It decides if the data transmission should take place on parallel paths or single path.
  2. It does the functions such as multiplexing, splitting or segmenting on the data.
  3. Transport layer guarantees transmission of data from one end to the other.
  4. It breaks the data groups into smaller units so that they are handled more efficiently by the network layer.

Layer 5 : The session layer : MU : May 15, Dec. 16

University Questions

Q. 1 Write short note on functions of Session layer. **(May 15, 5 Marks)**

Q. 2 Explain any four functions of session layer with example. **(Dec. 16, 10 Marks)**

- This layer manages and synchronizes conversations between two different applications.
- This is the level at which the user will establish system to system connection.
- It controls logging on and off, user identification, billing and session management.
- In the transmission of data from one system to the other, at session layer streams of data are marked and resynchronized properly so that the ends of messages are not cut prematurely and data loss is avoided.
- The session layer performs the following functions :
- It allows two systems to start a dialog. The communication between two processes will take place either in half duplex or full duplex mode.
- The other function of this layer is synchronization.
- The session layer is not inherently concerned with security and the network logon process.
- The primary functions of this layer is exchange of messages between two interested systems called as a **dialog.**
- In fact 22 different services are provided by the session layer.
- These are grouped into subsets such as the Kernel Function Unit, the Basic Activity Subset and the Basic Synchronization Subset.
- However the two most important services provided by the session layer are :
  1. Dialog control and 2. Dialog separation

1. Dialog control :

- Dialog control is the means by which a sending and receiving systems initiate a dialog, exchange messages and finally end the dialog.

2. Dialog separation :

- It is a process of inserting a reference marker called as a checkpoint into the data stream travelling between the sending and receiving systems.
- This allows the checking of status of both the machines at the same point in time.
- This will avoid any possible confusion and collision situation.

Layer 6 : The presentation layer :

- The presentation layer makes it sure that the information is delivered in such a form that the receiving system will understand and use it.
- The form and syntax (language) of the two communicating systems can be different Example, one system is using the ASCII code for file transfer and the other one uses IBM's EBCDIC.
- Under such conditions the presentation layer provides the "translation" from ASCII to EBCDIC and vice versa.
- The presentation layer is related to the **syntax** and **semantics** of the information being exchanged between the interested systems.

- Some of the important responsibilities of the presentation layer are :

  1. Translation 2. Encryption 3. Compression.

1. Translation :

- The communication systems usually exchange the information in the form of strings of characters, numbers etc.
- This information needs to be changed into bit streams before transmission.
- This is essential because different systems use different encoding techniques. The presentation layer does the job of translation.
- The presentation layer at the sending end converts the information into a common format and the presentation layer at the receiving end will convert this common format into the one which is compatible to the receiver.

2. Encryption :

- For ensuring the security and privacy of the information that is being communicated, a process called data encryption is essential.
- Encryption is carried out at the sending end. In the encryption process, the sender transforms the original information to another form, and sends the transformed information.
- At the receiving end, an exactly opposite process called Decryption is carried out in which the received information is transformed back to its original form.
- Encryption and Decryption are carried out by the presentation layer.

3. Compression :

- The data compression technique is used for reducing the number of bits required to send an information.
- Data compression is essential for transmission of multimedia such as text, audio and video.

Layer 7 : Application layer :

- *Application layer is at the top of all as shown in Fig. 1.15.2.*
- *It provides different services such as manipulation of information in various ways, retransferring the files of information, distributing the results etc. to the user who is sitting above this layer.*
- The functions such as LOGIN, or password checking are also performed by the application layer.
- The application layer performs the following functions :

  1. The application layer allows the creation of a virtual terminal which is the software version of a physical terminal. The user can log on to the remote host due to this arrangement.
  2. The application layer provides File Transfer Access and Management (FTAM) which allows a user to access, retrieve, manage or control files in a remote computer.
  3. It creates a basis for forwarding and storage of e-mails.

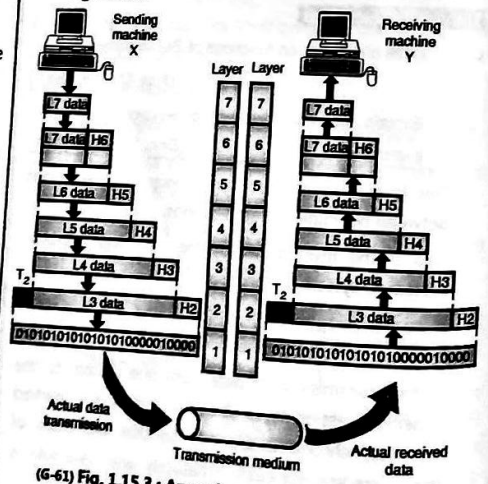
1.15.6 Exchange of Information in OSI Model :

MU : Dec. 09

University Questions

Q. 1 Explain the need for the layered architecture in computer network. Explain how information is exchanged between two nodes using OSI model. **(Dec. 09, 10 Marks)**

- At the physical layer, communication is direct i.e. machine X sends a stream of bits to machine Y.
- *At higher layers, each layer in the sending machine adds its own information to the message it receives from the layer just above it and passes the whole package to the layer just below it as shown in Fig. 1.15.3.*



(G-61) Fig. 1.15.3 : An exchange using the OSI model

- The information added by each layer is in the form of headers or trailers.
- Headers are added to the message at the layers 6, 5, 4, 3, and 2. A trailer is added at layer 2.
- At layer 1 the entire package is converted to a form that can be transferred to the receiving machine.
- At the receiving machine, the message is unwrapped layer by layer with each process receiving and removing the data meant for it.
- The upper OSI layers are always implemented in software (4, 5, 6 and 7) and lower layers are a combination of hardware and software (2, 3) except for the physical layer which is mostly hardware.
- Layers 1, 2 and 3 (i.e. physical , data link and network) are the network support layers.
- They deal with the physical aspects of moving data from one device to another such as electrical specifications, physical connections, physical addressing and transport timing and reliability.
- Layer 4, the transport layer ensures end to end reliable data transmission.
- Layers 5, 6 and 7 (i.e. session, presentation and application) they allow interoperability among unrelated software systems.

Protocols Associated With OSI reference Model :

Layer Name	Protocol Supported
Physical Layer	RS-232, RS-449, V.35, V.34, I.430, I.431, T1, E1, POTS, SONET/SDH, OTN, DSL, 802.11a/b/g/n, 802.15.x, USB · Bluetooth, Firewire (IEEE 1394)
Data Link Layer	ARP, CSLIP, SLIP, PPP
Network Layer	VLAN, IP (IPv4, IPv6), ICMP, IPsec, IGMP, IPX, AppleTalk
Transport Layer	TCP, UDP, SCTP, DCCP
Session Layer	NetBIOS, SAP, SIP, L2TP, VPN
Presentation Layer	MIME, XDR, TLS, SSL
Application Layer	NNTP = Usenet, SIP = VOIP, SSI, DNS, FTP, Gopher, HTTP, NFS, NTP = TIME, SMPP, SMTP = email, DHCP, SNMP, Telnet

1.15.7 Merits of OSI Reference Model :

- Some of the important advantages of OSI reference model are as follows :

  1. It distinguishes very clearly between the services, interfaces and protocols.
  2. The protocols in OSI model are better hidden. So they can be easily replaced by new protocols as the technology changes.
  3. OSI model is truly a general model.
  4. This model supports connection oriented as well as connectionless services.

1.15.8 Demerits of OSI Model :

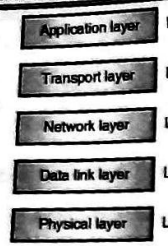
- Some of the important demerits of OSI reference model are as follows :

  1. Sessions and presentation layers are not of much use.
  2. This model was devised before the protocols were invented.

- So in real life there is a problem of fitting protocol into a model.

1.16 The TCP/IP Reference Model :

- TCP/IP is the short form of two important protocols namely Transmission Control Protocol/Internet Protocol.
- A **protocol suite** is defined as the set of protocols organized in different layers.
- The TCP/IP protocol suite is used in Internet today.
- TCP/IP is a hierarchical protocol suite means that each upper layer protocol receives support and services from either one or more lower level protocols.
- In the original TCP/IP protocol suite, there were four layers built upon the hardware.
- But today's TCP/IP protocol suite uses a five layer model as shown in Fig. 1.16.1.



(G-2065) **Fig. 1.16.1 : Layers in TCP/IP protocol suite**

1.16.1 Introduction to TCP/IP :

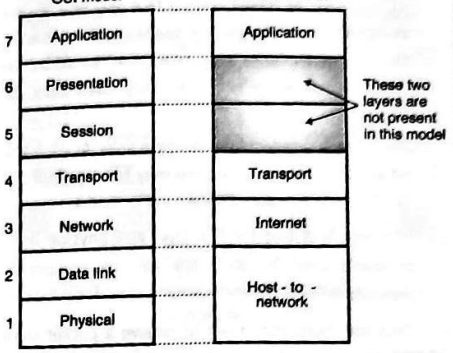
- The Internet protocol is like any other communication protocol is a set of rules which will govern every possible communication over the internet.
- Since the development of the ARPANET, TCP/IP together has emerged as the controlling body.
- It is being used in computers of not only in the U.S. but all over the world for all the types and sizes of computers.
- It has become the language of the Internet.
- TCP/IP are two protocols : Transmission control protocol and Internet protocol.
- These two protocols describe the movement of data between the host computers on Internet.
- The protocol however is a suite of many other protocols which provide for reliable communications across the Internet and the web.
- In the TCP/IP protocol suite, there are various layers, with each layer being responsible for different facets of communication.
- The Internet Protocol (IP) and Transmission Control Protocol (TCP) are together known as TCP/IP protocol.
- TCP/IP offers a simple naming and addressing scheme whereby different resources on Internet can be easily located.
- Information on Internet is carried in "packets". The IP protocol is used to put a message into a "packet".
- Each packet has the address of the sender and the recipient's address. These addresses are known as the IP addresses.
- Using the TCP protocol, a single large message is divided into a sequence of packets and each is put into an IP packet.
- The packets are passed from one network to another until they reach their destination.
- At the destination the TCP software reassembles the packets into a complete message.
- It is not necessary for all the packets in a single message to take the same route each time it is sent.

1.16.2 Layer Details of TCP/IP : MU : May 16

University Questions

Q. 1 Explain in short TCP/IP model. **(May 16, 4 Marks)**

- Transmission Control Protocol and the Internet Protocol (TCP/IP) was developed by the Department of Defence's Projects Research Agency (ARPA, later DARPA) under its project on network interconnection.
- It is a set of protocols that allow communication across multiple diverse network.
- ARPA originally created TCP/IP to connect military networks together, but later on this protocol was also given to government agencies and universities free of cost.
- Since the TCP/IP was developed for military use, it became robust to failures and flexible to diverse networks.
- TCP/IP is the most widely used protocol for interconnecting computers and it is the protocol of the Internet.
- TCP/IP became the standard for interoperating Unix Computers, especially in military and university environments.
- With the development of the Hypertext Transfer Protocol (HTTP) for sharing Hypertext Markup Language (HTML) documents freely on the internet, the World Wide Web (WWW) was born and soon TCP/IP came into much use.
- Fig. 1.16.2 shows the TCP/IP reference model along with the OSI model used for comparison.



(G-71) **Fig. 1.16.2 : TCP/IP reference model**

1.16.3 Description of TCP/IP Model :

MU : Dec. 12, May 16

University Questions

Q. 1 Explain the layer details of OSI and TCP/IP models. **(Dec. 12, 10 Marks)**

Q. 2 Explain in short TCP/IP model. **(May 16, 4 Marks)**

- As shown in Fig. 1.16.2, the TCP/IP model has only four layers.

Internet layer :

- This layer is called as the internet layer and it holds the whole architecture together.
- The task of this layer is to allow the host to insert packets into any network and then make them travel independently to the destination.
- The order in which the packets are received can be different from the sequence in which they were sent.
- Then the higher layers are supposed to arrange them in the proper order.
- Note that "Internet" is being used as a generic term.
- The internet layer defines (specifies) a packet format and a protocol called internet protocol (IP).
- The internet layer is supposed to deliver IP packets to their destinations.
- So routing of packets and congestion control are important issues related to this layer.
- Hence TCP/IP internet layer is very similar to the network layer in OSI model as shown in Fig. 1.16.2.

Transport layer :

- This is the layer above the internet layer. Its functions are same as those of a transport layer in OSI layer.
- This layer allows the peer entities of the source and destination machines to converse with each other.
- The end to end protocols used here are TCP and UDP (User datagram protocol).
- TCP is a reliable connection oriented protocol. It allows a byte stream transmitted from one machine to be delivered to the other machine without introducing any errors. TCP also handles the flow control.
- UDP (User datagram protocol) is the second protocol used in the transport layer.
- It is an unreliable, connectionless protocol and used for the applications which do not want the TCP's sequencing or flow control.
- UDP is also preferred over TCP in those applications in which prompt delivery is more important than accurate delivery. It is used in transmitting speech or video.

Application layer :

- TCP/IP model does not have session or presentation layers, because they are of little importance in most applications.
- The layer on top of transport layer is called as application layer.
- The protocols related to this layer are all high level protocols such as virtual terminal (TELNET), file transfer (FTP) and electronic mail (SMTP) as shown in Fig. 1.16.3.

Application layer	Telnet. FTP. SMIP. DNS. HTTP NNTP
Transport	TCP UDP
Internet (network)	IP
Host-to-network	Arpanet, satnet lan, paket radio

(G-2706) **Fig. 1.16.3**

- Many other protocols have been added to these, over the years such as Domain Name Service (DNS), NNTP and HTTP etc.

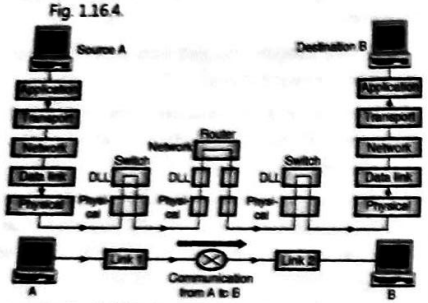
Host-to-network layer :

- This is the lowest layer in TCP/IP reference model.

- The host has to connect to the network using some protocol, so that it can send the IP packets over it.
- This protocol varies from host to host and network to network.

1.16.4 Layered Architecture :

- In order to understand how the communication takes place between various layers of TCP / IP protocol suite, we have considered a small internetwork consisting of three LANs (links) with all LANs connected to each other via a router as shown in Fig. 1.16.4.



(G-2178) **Fig. 1.16.4 : Communication through an Internet**

- In Fig. 1.16.4, there are two computers A and B communicating with each other and three more devices namely : the link layer switch in link-1, the router and the link layer switch in link-2.
- Computer A is called as the **source host** and computer B is called as the **destination host.**
- Each device in the Internet has a specific role to play, depending on which each device uses a set of layers as shown in Fig. 1.16.4.
- All the five layers are involved in communication for the source and destination hosts A and B respectively.
- At the source host, a message is created at the application layer and then it is sent in down the layers in order to physically send it to the destination host.
- At the destination host this message is received at the physical layer and then it is delivered to the application layer via the other layers between the physical and application layers.
- At the router, as shown in Fig. 1.16.4 only three layers of TCP/IP protocol suite are needed to be involved. Thus a router does not need the transport or application layers when it is being used only for routing.
- The router is connected to multiple links. At each link we use a switch which involves only two layers of the TCP/IP protocol suite as shown in Fig. 1.16.4.
- However note that the link layer and physical layer protocols used by each link can be completely different.
- Thus the router may have to receive a packet from link-1 based on one pair of protocol and may have to deliver a packet to link-2 based on a totally different pair of protocols.
- Now consider a switch in Fig. 1.16.4 which shows that it has two different connections. But both of them belong to the same link.
- Therefore two different protocol pairs will not be involved.
- A switch has to deal with only one pair of DLL and physical layer protocols.

1.16.5 Logical Connections in the TCP / IP:

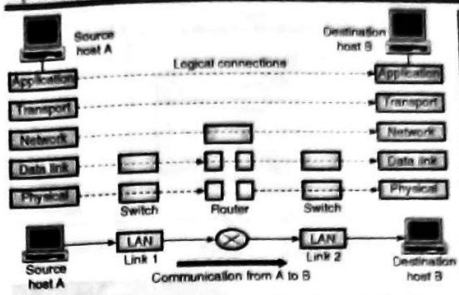
MU : Dec. 12, May 16

University Questions

Q. 1 Explain the layer details of OSI and TCP/IP models. **(Dec. 12, 10 Marks)**

Q. 2 Explain in short TCP/IP model. **(May 16, 4 Marks)**

- Now we are going to discuss the functions and duties of various layers in the TCP/IP protocol suite.
- In this section, we will think about the logical connections between various layers, so as to clearly understand the duties of each layer.
- The logical connections in a simple internetwork have been shown in Fig. 1.16.5.
- Each layer has some specific duties and we can use the logical connections to think about them easily.
- From Fig. 1.16.5 it is clear that the network, transport and application layers have an **end-to-end** duty.



(G-2177) **Fig. 1.16.5 : Logical connections between the layers of TCP / IP suite**

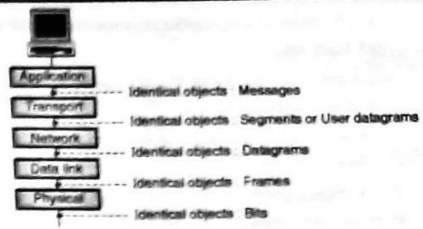
- But the data link and physical layers have the **hop to hop** duty. (Hop is a host or router).
- In this way the upper three layers have a **domain of duty** of the entire Internet while the lower two have a domain of duty of only link.

1.16.6 Data unit created by every layer :

- We can think about the logical connections in a different way i.e. in terms of the **data unit** created by each layer.
- The names of data units (packets) created by different layers are as follows :

Layer	Data unit	Layer	Data unit
Application	Message	Datalink	Frame
Transport	Segment	Physical	Bits
Network	Datagram		

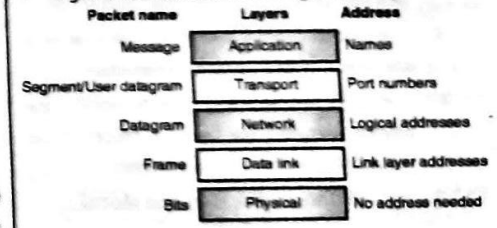
- The data unit (packet) created by the top three layers, should not be changed by a router or a link layer switch.
- However the data unit created at the lower two levels can be changed only by the router. The link layer switches cannot modify it.
- The second principle that we discussed for the protocol layering has been shown in Fig. 1.16.6.
- Note that the objects shown below each layer related to each device are identical.



(G-2179) **Fig. 1.16.6 : Identical objects in the TCP/IP suite**

1.17 Addressing in TCP/IP :

- Addressing is another important concept related to the protocol layering in the Internet.
- There is a logical connection between the pair of layers as discussed earlier.
- For any communication to take place between a source and a destination, two addresses namely source address and destination address are needed.
- Thus we will need four pairs of such addresses corresponding to the data link, network, transport and application layers.
- There is no need of addresses at the physical layer because communication at the physical layer takes place in bits which can not have an address.
- Fig. 1.17.1 shows the addressing at each layer.



(G-2069) **Fig. 1.17.1 : Addressing in TCP/IP protocol suite**

- Fig. 1.17.1 also shows the relationship between various layers, the addresses used in each layer and the name of the packet at each layer.
- We generally use the **names** to define the **site address** which provides the required services. For example **techknowledgebook.com,** at the application layer. It is also possible to use the email address such as **jayantkatre@gmail.com.**

- The addresses at the transport layer are called as **port numbers.**
- These define the programs at the application layer of source and destination.
- There are several application layer programs running at a time.
- Port numbers are the local addresses which are used to distinguish between these programs.
- The addresses at the network layer are global in nature because the whole Internet is the scope of these addresses.
- The connection of a device to the Internet is uniquely defined by a network layer address.
- The addresses at the data link layer are called as the **MAC addresses.**
- These are the locally defined addresses. Each host or router in a network such as LAN or WAN always has a MAC address.

1.17.1 Demerits of TCP/IP Model :

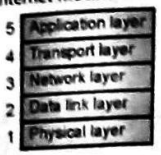
1. TCP/IP model does not clearly distinguish the concepts of service, interface and protocol.
2. This model is not at all general and it cannot describe any protocol stack other than TCP/IP.
3. The host-to-network layer is not a layer at all in the normal sense. It is simply an interface.
4. The TCP/IP model does not even mention the physical and data link layers. A proper model should include both as separate layers.

1.17.2 Hybrid (Internet) Reference Model :

- In spite of many problems associated with the OSI model, it has proved to be very useful one practically.
- But the OSI protocols have not become popular.
- On the other hand the TCP/IP model is practically non existing but the TCP/IP protocols are used widely.

So sometimes a modified OSI model with primary concentration on TCP/IP is used which is called as the hybrid model.

- The hybrid model is shown in Fig. 1.17.2. It is also called as the Internet model.



(G-2704) **Fig. 1.17.2 : Hybrid model**

1.17.3 Comparison of Models :

MU : Dec. 14, May 17, Dec. 17

University Questions

Q. 1 Why there is a need for layered designing for networking and communication ? Compare the TCP/IP and OSI reference models. **(Dec. 14, 10 Marks)**

Q. 2 What is ISO-OSI reference model ? Compare it with TCP/IP reference model. Which layer is used for the following :

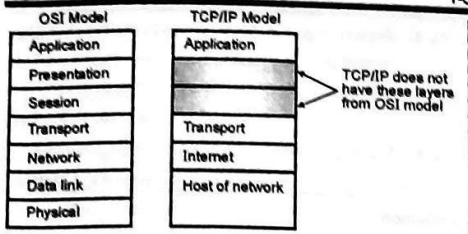
1. To route packets
2. To convert packets to frame
3. To detect and correct errors.
4. To run services like FTP, telnet etc.

(May 17, 10 Marks)

Q. 3 Explain the need of layered design for communication and networking. Compare the OSI reference model & TCP/IP. **(Dec. 17, 10 Marks)**

Similarities between OSI and TCP/IP Models :

- Following are some of the similarities between OSI and TCP/IP models :
  1. In both the models the functions of layers is approximately same.
  2. Both models use the concept of layered architecture.
  3. The transport layers and the layers below it provide transport services independent of networks.
  4. In both the models, the layers above transport layer are application oriented.
- Refer to Fig. 1.17.3 and Table 1.17.1 for the comparison of the two reference models.



(G-73) **Fig. 1.17.3 : Relationship between OSI and TCP/IP models**

Differences between OSI and TCP/IP Models :

- The differences between these two reference models are as given below.

Table 1.17.1 : Difference between OSI and TCP/IP model

OSI	TCP/IP
Has 7 layers	Has 4 layers
Transport layer guarantees delivery of packets.	Transport layer does not guarantee delivery of packets.
Horizontal approach.	Vertical approach.
Separate session layer.	No session layer, characteristics are provided by transport layer.
Separate presentation layer.	No presentation layer, characteristics are provided by application layer.
Network layer provides both connectionless and connection oriented services.	Network layer provides only connection less services.
It defines the services, interfaces and protocols very clearly and makes a clear distinction between them.	It does not clearly distinguish between service, interfaces and protocols.
The protocols are better hidden and can be easily replaced as the technology changes.	It is not easy to replace the protocols.
OSI is truly a general model.	TCP/IP cannot be used for any other application.
It has a problem of protocol fitting into a model.	The model does not fit any other protocol stack.

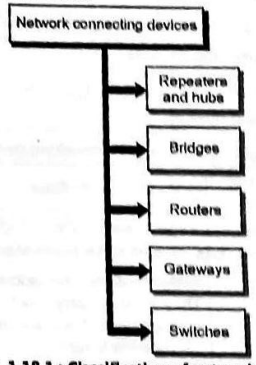
1.18 Networking Devices :

MU : Dec. 15, Dec. 18

University Questions

Q. 1 Write short notes on : Internetworking devices. **(Dec. 15, Dec. 18, 10 Marks)**

- Different types of network connecting devices are as shown in Fig. 1.18.1.

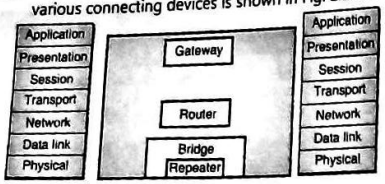


(G-348) **Fig. 1.18.1 : Classification of networking devices**

Networking devices :

- Two or more devices are connected to each other for the purpose of sharing data or resources from a network.
- A LAN may be spread over a larger distance than its media can handle effectively.
- The number of stations also can be more than a number which can be handled and managed properly.
- Such networks should be subdivided into smaller networks and these smaller subnetworks should be connected to each other through connecting devices.
- A device called a **repeater** is inserted into the network to increase the coverable distance or a device called a **bridge** can be inserted for traffic management.
- When two or more separate networks are connected for exchanging data or resources it creates an internetwork.
- Routers and gateways are used for internetworking. Each of these device type interacts with protocols at different layers of the OSI model.

- The relation between OSI reference model and various connecting devices is shown in Fig. 1.18.2.



(G-806(a)) **Fig. 1.18.2 : Connecting devices and OSI model**

Roles of connecting devices :

- Table 1.18.1 summarizes the roles of different networking devices.

Table 1.18.1 : Roles of networking devices

Sr. No.	Name of the device	Role
1.	Repeater	Regenerates the original signal. Operates in the physical layer.
2.	Bridge	Bridges utilize the address protocol. They can carry out the traffic management. They are most active in the data link layer.
3.	Routers	Routers provide connections between two separate but compatible networks. It works in the network layer.
4.	Gateways	Gateways provide translation services between incompatible networks and works in all the layers.
5.	HUB	Connecting stations in a physical layer topology.
6.	Switch	Provides bridging functionality with great efficiency.

1.19 Hubs :

MU : May 10, May 11, Dec. 11, May 12, Dec. 12, Dec. 13, Dec. 15, May 19

University Questions

Q. 1 Explain working of following network components and state in which layer they work. Repeaters, Hubs, Bridges, Routers, Switches, Gateways. **(May 10, May 11, Dec. 11, 10 Marks)**

Q. 2 What a neat diagram compare the uses and functions of different hardware components/devices used in an internetwork. **(May 12, 10 Marks)**

Q. 3 Explain with example : Hubs. **(Dec. 12, 2 Marks)**

Q. 4 Explain the functions of the different network hardware components. **(Dec. 13, 10 Marks)**

Q. 5 Write short notes on : Internetworking devices. **(Dec. 15, 10 Marks)**

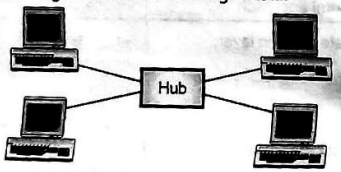
Q. 6 Explain Repeater, Hub, Bridge, Switch, Gateway. **(May 19, 5 Marks)**

Definition :

- The general meaning of the word hub is any connecting device. But its specific meaning is **multiport repeater.**

Function :

- It is normally used for connecting stations in a physical star topology.
- All networks require a central location to connect various segments of media coming from various nodes.
- Such a central location is called as a hub. A hub organizes the cables and relays signals to the other media segments as shown in Fig. 1.19.1.



(G-350) ***Fig. 1.19.1 : Hub***

Types :

- There are three main types of hubs :
  1. Passive hubs
  2. Active hubs
  3. Intelligent hubs

1.19.1 Passive Hubs :

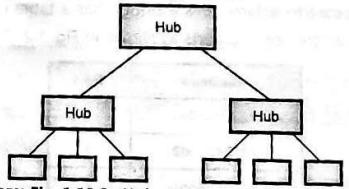
- A passive hub simply combines the signals of a network segments.
- There is no signal processing or regeneration. It merely acts as a connector.
- A passive hub reduces the cabling distance by half because it does not boost the signals and in fact absorbs some of the signal.
- With a passive hub, each computer receives the signals sent from all the other computers connected to the hub.

1.19.2 Active Hubs :

- They are like passive hubs but have electronic components for regeneration and amplification of signals.
- By using active hubs the distance between devices can be increased. An active hub is equivalent to a multipoint repeater.
- The main drawback of active hubs is that they amplify noise as well along with the signals. They are more expensive than passive hubs as well.

1.19.3 Intelligent Hubs :

- In addition to signal regeneration, intelligent hubs perform some other intelligent functions such as network management and intelligent path selection.
- A switching hub chooses only the port of the device where the signal needs to go, rather than sending the signal along all paths.
- Hubs can also be used to create multiple levels of hierarchy as shown in Fig. 1.19.2.



(G-353) **Fig. 1.19.2 : Hubs to create multiple levels of hierarchy**

1.20 Repeaters :

MU : May 10, May 11, Dec. 11, May 12, Dec. 12, Dec. 13, Dec. 15, May 19

University Questions

Q. 1 Explain working of following network components and state in which layer they work. Repeaters, Hubs, Bridgers, Routers, Switches, Gateways. **(May 10, May 11, Dec. 11, 10 Marks)**

Q. 2 What a neat diagram compare the uses and functions of different hardware components/devices used in an internetwork. **(May 12, 10 Marks)**

Q. 3 Explain with example : Repeater. **(Dec. 12, 2 Marks)**

Q. 4 Explain the functions of the different network hardware components. **(Dec. 13, 10 Marks)**

Q. 5 Write short notes on : Internetworking devices. **(Dec. 15, 10 Marks)**

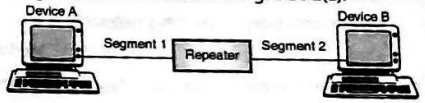
Q. 6 Explain Repeater, Hub, Bridge, Switch, Gateway. **(May 19, 5 Marks)**

Definition :

- A repeater is a connecting device which can operate only in the physical layer of the OSI model.
- All transmission media weaken the electromagnetic waves that travel through them.
- Attenuation of signals limits the distance any medium can carry data. Devices that amplifies signals to ensure data transmission are called **repeaters.**

Function of a repeater :

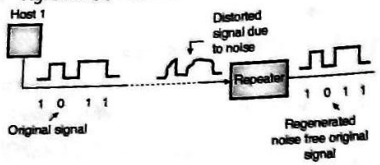
- A repeater receives a signal and before it gets attenuated or corrupted, regenerates the original signal.
- Thus we can use a repeater to extend the physical length of LAN as a shown in Fig. 1.20.1(a).



(G-351) **Fig. 1.20.1(a) : Repeater**

- Repeater is not an amplifier because amplifiers simply amplify the entire incoming signal along with noise.
- Signal– regenerating repeaters create an exact duplicate of incoming data by identifying it amidst the noise, reconstructing it and retransmitting only the desired information.
- The original signal is duplicated, boosted to its original strength and sent as shown in Figs. 1.20.1(a) and (b).
- A repeater does not connect two LANs. It connects only two devices connected in the same LAN.
- It cannot connect two LANs of a different protocols.
- A repeater forwards every frame, it cannot filter out some frames and let the others pass through.
- A repeater should be placed at a precise point on the link such that the signal reaches it before the noise has induced an error in any of the transmitted bits.

- Fig. 1.20.1(b) illustrates the function of a repeater.



(G-352) **Fig. 1.20.1(b) : Function of a repeater**

- Repeaters operate at the physical layer of the OSI model and they deal with the actual physical signals.

1.21 Bridges :

MU : May 10, May 11, Dec. 11, May 12, Dec. 12, Dec. 13, Dec. 15

University Questions

Q. 1 Explain working of following network components and state in which layer they work. Repeaters, Hubs, Bridges, Routers, Switches, Gateways. **(May 10, May 11, Dec. 11, 10 Marks)**

Q. 2 What a neat diagram compare the uses and functions of different hardware components/devices used in an internetwork. **(May 12, 10 Marks)**

Q. 3 Explain with example : Bridges. **(Dec. 12, 2 Marks)**

Q. 4 Explain the functions of the different network hardware components. **(Dec. 13, 10 Marks)**

Q. 5 Write short notes on : Internetworking devices. **(Dec. 15, 10 Marks)**

Definition :

- Bridge is a computer network device, which creates a single aggregate network from multiple communication networks or network segments. This function is called as **network bridging.**

Function :

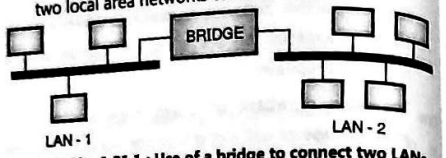
- A bridge can operate in the physical layer as well as in the data link layer of the OSI model.
- It can regenerate the signal that it receives and it can check the physical (MAC) addresses of source and destination mentioned in the header of a frame.

Types of bridges :

- There are four types of bridges :
  1. Simple bridges.
  2. Multipoint bridges.
  3. Transparent bridges.
  4. Source routing bridges.

Configuration :

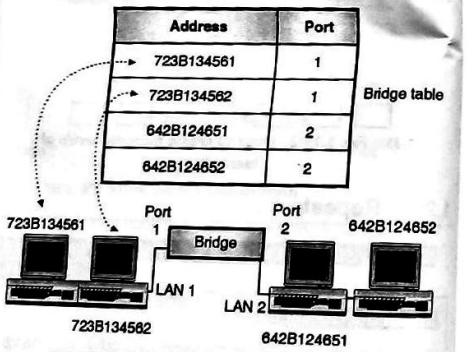
- Fig. 1.21.1 shows the use of a bridge for connecting two local area networks to each other.



(G-2377) **Fig. 1.21.1 : Use of a bridge to connect two LANs**

Operation :

- The major difference between the bridge and repeater is that the bridge has a filtering capability.
- That means a bridge will check the destination address of a frame and make a decision about whether the frame should be forwarded or dropped.
- If the frame is to be forwarded, then the bridge should specify the port over which it should be forwarded.
- In order to achieve this a bridge has a table relating the addresses and ports as shown in Fig. 1.21.2.



(G-354) **Fig. 1.21.2 : Bridge and bridge table**

- If a frame for 723B134561 arrives at port 2 then the bridge goes through its table and understands that the frame is to be sent out on port 1 so it will do so.
- In Fig. 1.21.2 a two port bridge is shown but in reality a bridge has more than two ports.
- It is important to note that the bridges do not change the physical address contained in the frame.

Uses of bridges :

1. A bridge joins two otherwise separate computer networks.
2. Bridges are used with LANs to extend their reach to cover larger physical area.
3. Bridges are used to inspect the incoming network traffic and determine whether to forward it or discard it.

1.22 Routers :

MU : May 10, May 11, Dec. 11, May 12, Dec. 12, Dec. 13

University Questions

Q. 1 Explain working of following network components and state in which layer they work. Repeaters, Hubs, Bridges, Routers, Switches, Gateways. **(May 10, May 11, Dec. 11, 10 Marks)**

Q. 2 What a neat diagram compare the uses and functions of different hardware components/devices used in an internetwork. **(May 12, 10 Marks)**

Q. 3 Explain with example : Router. **(Dec. 12, 2 Marks)**

Q. 4 Explain the functions of the different network hardware components. **(Dec. 13, 10 Marks)**

Definition :

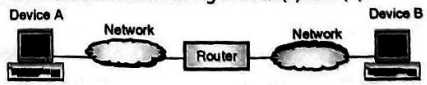
- Router is a networking device which forwards the data packets between computer networks.
- Routers perform the traffic directing function on the Internet.

Functions :

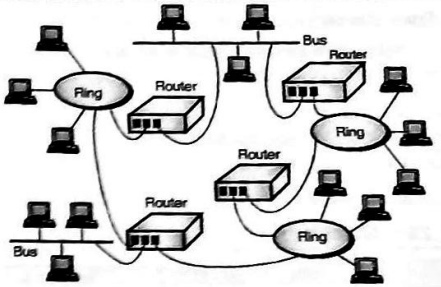
- The two important functions, performed by a router are :
  1. Determination of path (routing)
  2. Packet forwarding.

OSI layer :

- Routers work at the network layer of the OSI model.
- Routers are devices that connect two or more networks as shown in Figs. 1.22.1(a) and (b).



(G-364) **Fig. 1.22.1(a) : A router**



(G-365) **Fig. 1.22.1(b) : Routers in an internet**

- They consist of a combination of hardware and software.
- The hardware can be in the form of a network server, a separate computer or a special device, as well as the physical interfaces to the various networks in the internetwork.
- Various types of networks can be interconnected through routers as shown in Fig. 1.22.1(b).
- The software in a router are the operating system and the routing protocol. Management software can also be used.
- Routers use logical and physical addressing to connect two or more logically separate networks.
- The large network is organized into small network segments called as subnets and these subnets are interconnected via routers.
- Each of the subnet is given a logical address. This allows the networks to be separate but still access each other and exchange data.
- Data is grouped into packets, or blocks of data. Each packet has a physical device address as well as logical network address.
- The network address allows routers to calculate the optimal path to a workstation or computer.
- Route discovery is the process of finding the possible routes through the internetwork and then building routing tables to store that information.
- The two methods of route discovery are :
  1. Distance vector routing
  2. Link state routing.

Types of routers :

- Following are different types of routers :
  1. Wired routers.
  2. Wireless routers.
  3. Core and edge routers.
  4. Virtual routers.
  5. Broadband routers.

1.23 Gateways :

MU : May 11, Dec. 11, May 12, Dec. 13, May 19

University Questions

Q. 1 Explain working of following network components and state in which layer they work. Repeaters, Hubs, Bridgers, Routers, Switches, Gateways. **(May 11, Dec. 11, 10 Marks)**

Q. 2 What a neat diagram compare the uses and functions of different hardware components/devices used in an internetwork. **(May 12, 10 Marks)**

Q. 3 Explain the functions of the different network hardware components. **(Dec. 13, 10 Marks)**

Q. 4 Explain Repeater, Hub, Bridge, Switch, Gateway. **(May 19, 5 Marks)**

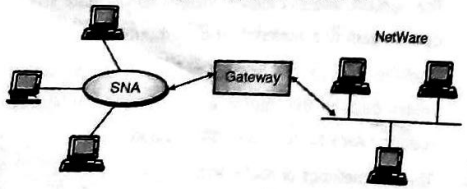
Definition :

- *Gateway is a network connecting device that connects two dissimilar networks together.*

Diagram :

- Fig 1.23.1(b) shows the diagram showing a gateway connecting two dissimilar networks SNA and LAN.

(a) A gateway in the OSI model



(b) A gateway

(G-366) Fig. 1.23.1

When to use a gateway ?

- When the networks that must be connected are using completely different protocols from each other, a powerful and intelligent device called a **gateway** is used.

Functions of a gateway :

- A gateway is a device that can interpret and translate different protocols that are used on two distinct networks as shown in Figs. 1.23.1(a) and (b).
- Gateways comprise of software, dedicated hardware or a combination of both. Gateway operate through all the seven layers of the OSI model and all five layers of the internet model.
- A gateway can actually convert data so that it works with an application on a computer on the other side of the gateway.
- For e.g. a gateway can receive e-mail message in one format and convert them into another format.
- Gateways can connect systems with different communication protocols, languages and architecture.
- For e.g. IBM networks using Systems Network Architecture (SNA) can be connected to LANs using a gateway.

Note : *Gateways are slow because they need to perform intensive conversions.*

1.24 Switches :

MU : May 11, Dec. 11, May 12, Dec. 12, Dec. 13, May 19

University Questions

Q. 1 Explain working of following network components and state in which layer they work. Repeaters, Hubs, Bridgers, Routers, Switches, Gateways. **(May 10, May 11, Dec. 11, 10 Marks)**

Q. 2 What a neat diagram compare the uses and functions of different hardware components/ devices used in an internetwork. **(May 12, 10 Marks)**

Q. 3 Explain with example : Switches. **(Dec. 12, 2 Marks)**

Q. 4 Explain the functions of the different network hardware components. **(Dec. 13, 10 Marks)**

Q. 5 Explain Repeater, Hub, Bridge, Switch, Gateway. **(May 19, 5 Marks)**

Definition :

- A switch is a networking device which connects devices together on a computer network by using packet switching to receive, process and forward data to the destination.

Which layer of OSI model ?

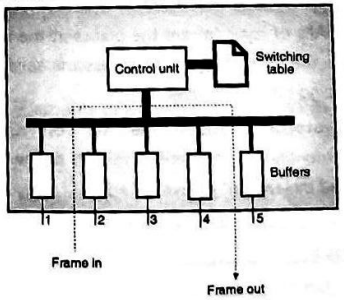
- Switches generally operate at the data link layer of the OSI model but some switches can work at the network layer too.

Role of a switch :

- A switch is a device which provides bridging functionality with greater efficiency.
- A switch acts as a multiport bridge to connect devices or segments in a LAN.
- The switch has a buffer for each link to which it is connected.
- When it receives a packet, it stores the packet in the buffer of the receiving link and checks the address to find the outgoing link.
- If the outgoing link is free, the switch sends the frame to that particular link.

Types of switches :

- Switches are of two types :
  1. Store - and - forward switch
  2. Cut - through switch.
- A store - and - forward switch stores the frame in the input buffer until the whole packet has arrived.
- A cut-through switch, forwards the packet to the output buffer as soon as the destination address is received.
- Concept of a switch is shown in Fig. 1.24.1. As shown in the Fig. 1.24.1 a frame arrives at port 2 and is stored in the buffer.



(G-367) Fig. 1.24.1 : Switch

- The CPU and the control unit, using the information in the frame consult the switching table to find the output port. The frame is then sent to port 5 for transmission.

Note : Routing switches use the network layer destination address to find the output link to which the packet should be forwarded.

1.24.1 Comparison of Networking Devices:

MU : Dec. 09, Dec. 10

University Questions

Q. 1 Differentiate between Hub and Switch. **(Dec. 09, Dec. 10, 6 Marks)**

Sr. No.	Parameter	Hub	Bridge	Switch	Router	Gateways
1.	Layer in OSI model	Physical	Physical and datalink	Datalink	Network	Seven layers
2.	Operation	Organises the cables and relays signals to other media segments	Regeneration, checks MAC address	Switch is a multiport bridge to connect devices or segments in a LAN	Connects two or more networks	Provides translation services between in-compatible networks
3.	Types	Passive, Active and Intelligent	Transparent and Routing	Two layer and Three layer	Distance vector and Link state	-
4.	Cost	Low cost	Very expensive	Expensive	Low cost	Costly

Review Questions

- **Q. 1** State various services provided by the network for companies and people.
- **Q. 2** What is the difference between broadcast and point to point networks ?
- **Q. 3** What is meant by internetwork ?
- **Q. 4** Write a short note on MAN.
- **Q. 5** Write a short note on WAN.
- **Q. 6** Compare LAN, WAN and MAN.
- **Q. 7** Define peer.
- **Q. 8** How does the actual data transfer take place between two machines.

2.1 Introduction to Communication:

Definition :

- Communication is defined as the act of transferring information from one place, person or group to another.
- Every communication involves (at least) one sender, a message and a recipient.
- The communication branch is the oldest branch of the electronics field. Telecommunication means communicating at a distance.
- A communication system is the means of conveying the information from one place to the other.

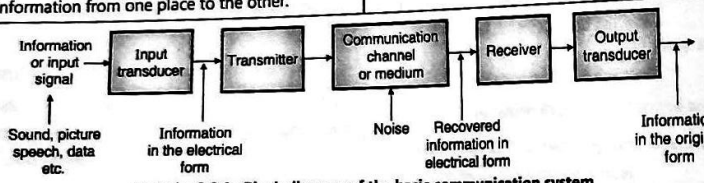
2.2 Elements of Communication System :

Definition :

- A communication system is defined as a collection of individual communications networks, transmission systems, relay stations, communication channels and receivers interconnected to exchange meaningful information.

Block diagram :

- The block diagram of the simplest possible communication system is as shown in Fig. 2.2.1.



(D-1) Fig. 2.2.1 : Block diagram of the basic communication system

- As seen from the Fig. 2.2.1, the elements of a basic communication system are transmitter, a communication medium (channel) and the receiver.
- When the transmitted signal is travelling from the transmitter to the receiver over a communication channel, noise gets added to it.
- The elements of basic communication system are as follows :
  1. Information or input signal
  2. Input transducer
  3. Transmitter
  4. Communication channel or medium
  5. Noise
  6. Receiver
  7. Output transducer

Information or input signal :

- The communication systems have been developed for communicating useful information from one place to the other.
- This information can be in the form of a sound signal like speech or music, or it can be in the form of pictures (TV signals) or it can be data information coming from a computer.

Input transducer :

- The information in the form of sound, picture or data signals cannot be transmitted as it is.
- First it has to be converted into a suitable electrical signal. The input transducer block does this job.
- The input transducers commonly used in the communication systems are microphones, TV camera etc.

Transmitter :

- The function of the transmitter block is to convert the electrical equivalent of the information to a suitable form.
- In addition to that it increases the power level of the signal.
- The power level should be increased in order to increase the range of transmitted signal.
- The transmitter consists of the electronic circuits such as amplifier, mixer, oscillator and power amplifier.

Communication channel or medium :

- The communication channel is the path used for transmission of electronic signal from one place to the other.

- The communication medium can be conducting wires, cables, optical fibre or free space.
- Depending on the type of communication medium, two types of communication systems will exist.
- They are :
  1. Wired communication or line communication
  2. Wireless communication or radio communication

1. Line communication :

- The line communication systems use the communication mediums like the simple wires or cables or optical fibers.
- The examples of such systems, are telegraph and telephone systems, cable T.V. etc.
- Due to physical connection from one point to the other, these systems cannot be used for the communication over long distances.

2. Radio communication :

- The radio communication systems use the free space as their communication medium.
- They do not need the wires for sending the information from one place to the other.
- The radio or TV broadcasting, satellite communication are the examples of the wireless communication. These systems transmit the signal using a transmitting antenna in the free space.
- The transmitted signal is in the form of electromagnetic waves.
- A receiving antenna will pick up this signal and feed it to the receiver.
- Radio communication can be used for the long distance communication such as from one country to the other or even from one planet to the other.

Noise :

- Noise is an unwanted electrical signal which gets added to the transmitted signal when it is travelling towards the receiver.
- Due to noise, the quality of the transmitted information will degrade.
- Once added, the noise cannot be separated out from the information.
- Hence noise is a big problem in the communication systems. (Especially analog communication systems).
- The noise can be either natural or manmade. The sources of natural noise are lightning or radiation from the sun and stars etc.
- The man made noise includes the noise produced by electrical ignition systems of the automobiles, welding machines, electric motors etc.
- Even though noise cannot be completely eliminated, its effect can be reduced by using various techniques.

Receiver :

- The process of reception is exactly the opposite process of transmission.
- The received signal is amplified, demodulated and converted into a suitable form.
- The receiver consists of electronic circuits like mixer, oscillator, detector, amplifier etc.
- Various types of receivers are used depending on the requirements of applications.

Output transducers :

- The output transducer converts the electrical signal at the output of the receiver back to the original form i.e. sound or TV pictures etc.
- The typical examples of the output transducers are loud speakers, picture tubes, computer monitor etc.

2.3 The Electromagnetic Spectrum :

Electromagnetic waves :

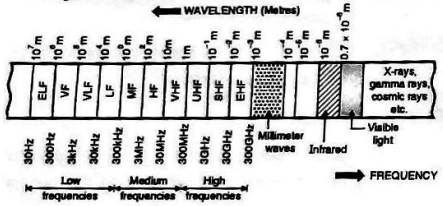
- The information signal should be first converted into an electromagnetic signal before transmission because the wireless transmission takes place using the electromagnetic waves.
- The electromagnetic waves consist of both electric and magnetic fields.
- The electromagnetic waves can travel a long distance through space.
- The electromagnetic signals are also called as radio frequency (RF) waves.
- The EM waves oscillate. They are sinusoidal and their frequency is measured in Hz.

Definition :

- The frequency of EM signal can be very low or it can be extremely high.
- This entire range of frequencies of EM waves is called as **Electromagnetic spectrum.**

Description :

- The electromagnetic spectrum consists of signals such as 50 Hz line frequency and voice signals at the lower end.
- The radio frequencies which are used for the two way communication reside at the center of the EM spectrum.
- These frequencies are used for the applications such as radio or TV broadcasting as well.
- The infrared and visible light are at the upper end of the EM spectrum.
- Fig. 2.3.1 shows the entire electromagnetic spectrum.



(D-16)Fig. 2.3.1 : Complete electromagnetic (EM) spectrum

- *The short forms used in the EM spectrum of Fig. 2.3.1 have the following meanings.*

Table 2.3.1 : Segments of the electromagnetic spectrum

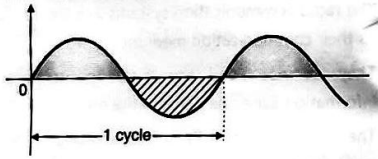
Sr. No.	Name	Frequency	Wavelength
1.	Extremely low frequencies (ELF)	30-300 Hz	10^7 to 10^6 m
2.	Voice frequencies (VF)	300-3000 Hz	10^6 to 10^5 m
3.	Very low frequencies (VLF)	3-30 kHz	10^5 to 10^4 m
4.	Low frequencies (LF)	30-300 kHz	10^4 to 10^3 m
5.	Medium frequencies (MF)	300 kHz - 3 MHz	10^3 to 10^2 m
6.	High frequencies (HF)	3-30 MHz	10^2 to 10 m
7.	Very high frequencies (VHF)	30-300 MHz	10 to 1 m
8.	Ultra high frequencies (UHF)	300 MHz-3 GHz	1 to 10^{-1} m
9.	Super high frequencies (SHF)	3-30 GHz	10^{-1} to 10^{-2} m
10.	Extremely high frequencies (EHF)	30-300 GHz	10^{-2} to 10^{-3} m
11.	Infrared	30 to 430 THz	0.7 to 10 μm
12.	Visible light	375-750 THz	0.4 μm to 0.8 μm

2.3.1 Frequency and Wavelength :

- In the EM spectrum, we have used frequency as well as wavelength in order to define various segments.
- So let us define these terms and the relation between them.

Frequency :

- *Frequency is defined as the number of cycles of a waveform per second. It it expressed in hertz (Hz).*
- Refer Fig. 2.3.2 for the concept of frequency.



(D-17) Fig. 2.3.2 : One cycle

- The units used to measure higher frequencies are kilohertz (kHz), Megahertz (MHz) and Gigahertz (GHz). Their relation with the basic unit Hz is as follows :

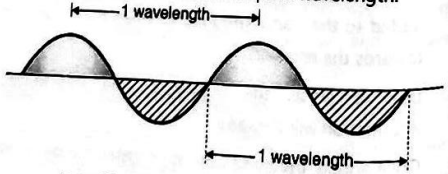
1 kHz = 1000 Hz

1 MHz = 1000 kHz = 1×10^6 Hz.

1 GHz = 1000 MHz = 1×10^9 Hz.

Wavelength (λ) :

- Wavelength is defined as the distance travelled by an electromagnetic wave during the time of one cycle.
- Refer Fig. 2.3.3 for the concept of wavelength.



(D-18) Fig. 2.3.3 : Definition of wavelength

- Since EM waves travel at the speed of light in the free space or vacuum, their wavelength is given by,

$$\lambda = \frac{\text{Speed of light}}{\text{Frequency}} = \frac{3 \times 10^8 \text{ m/s}}{f} \quad \ldots(2.3.1)$$

- Hence wavelength decreases with increase in frequency.

2.3.2 EM Spectrum and Communication Applications :

- In the radio communication system the frequencies ranging from a few kilohertz to many gigahertz all are being used for various purposes.
- Different frequency ranges defined for communication applications are as given below :

1. Extremely low frequencies (ELF) :

- The frequencies in the range 30 Hz - 300 Hz are called as Extremely low frequencies (ELF).

2. Voice frequencies (VF):

- The frequencies in the range 300 Hz – 3 KHz are called as voice frequencies (VF).

3. Very low frequencies (VLF):

- The frequencies in the range 3 kHz to 30 kHz are called as very low frequencies (VLF).

4. Low frequencies (LF) :

- The frequencies in the range 30 kHz to 300 kHz are known as the low frequencies (LF).

5. Medium frequencies (MF) :

- The frequencies most commonly used in early days were from about 300 kHz to 3 MHz and were called as medium frequencies (MF).

6. High frequencies (HF) :

- On the higher frequency side high frequencies (HF) will cover the frequency range from 3 MHz to 30 MHz.

7. Very high frequency (VHF) :

- Then very high frequency (VHF) from 30 MHz to 300 MHz and so on.
- The frequencies beyond 300 MHz are classified into UHF, SHF and EHF bands
- Table 2.3.2 gives you the details of entire usable frequency spectrum and its applications.

Table 2.3.2 : The Radio Frequency Spectrum with applications

Sr. No.	Frequency band	Applications
1.	30 Hz - 300 Hz. Extremely low frequencies ELF.	Power transmission.
2.	300 Hz - 3 kHz. Voice frequencies (VF)	Audio applications.
3.	3 kHz – 30 kHz. Very low frequencies (VLF)	Submarine communications. Navy, Military communications
4.	30 kHz - 300 kHz. Low frequencies (LF)	Aeronautical and marine, navigation, these frequencies act as sub carriers.
5.	300 kHz - 30 MHz Medium frequencies (MF)	AM radio broadcast, Marine and aeronautical communications.
6.	3 MHz - 30 MHz. High frequencies (HF)	Short-wave transmission, Amateur and CB communication.
7.	30 MHz - 300 MHz Very high frequencies (VHF)	TV broadcasting, FM broadcasting.
8.	300 MHz - 3 GHz Ultra high frequencies (UHF)	UHF TV channels ,Cellular phones, Military applications
9.	3 GHz - 30 GHz (SHF)	Satellite communication and Radar
10.	30 - 300 GHz (EHF)	Satellites and specialized radars

Infrared Signals :

- The EM signals having frequencies above 300 GHz are not referred as radio waves.
- The signal occupying the range between 0.1 mm and 700 nanometers (nm) are called infrared signals.
- These are used in various special types of communications.
- Some of them are as follows :
  1. In astronomy to detect stars and other heavenly bodies.
  2. In the guided weapon systems.
  3. TV remote control.
  4. Wireless keyboards and mouse.

Visible Light :

- Light is a special type of electromagnetic radiation. It has wavelength in the range of 0.4 to 0.8 μm.
- Light is used for various kinds of communications.
- Light waves can be modulated using the signal to be transmitted and transmitted through the glass fibers in the optical fiber communication system.
- Light signals can also be transmitted through free space. Laser is a type of light, which can be easily modulated with voice, video and data information.

2.4 Introduction to Physical Layer :

- In this chapter we are going to discuss about the lowest layer in the OSI model called the physical layer.
- The physical layer defines the mechanical, electrical and timing interfaces to the network.
- This chapter also deals with the fundamental limits put by the nature on data transmission over a channel.
- Latter on in the chapter various types of transmission media such as guided, wireless and satellite have been discussed.

2.4.1 Physical Layer Design Issues :

- The major task of physical layer is to provide services for the data link layer.
- Following are the services provided or duties of the physical layer or the design issues of the physical layer.

(G-91) Fig. 2.4.1 : Physical layer design issues

Transformation from bit to signal :

- The data link layer consists of 0's and 1's in the bit form.
- This bit stream cannot travel as it is on the transmission medium.
- So the physical layer converts the bit stream into a signal which is suitable for the transmission medium.

Bit rate control :

- The highest value of bit rate depends on the transmission medium being used and the physical layer acts as a bit rate controller.
- The design of the physical layer hardware and software will determine the data rate.

Bit synchronization :

- The timing related to the data bit transfer is very important in computer communication.
- The physical layer governs the synchronization of the bits by providing a clock which controls the transmitter as well as receiver.

Multiplexing :

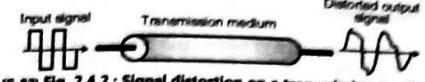
- Physical layer can use different techniques of multiplexing, in order to improve the channel efficiency.

Switching :

- There are three switching methods, namely circuit switching, message switching and packet switching.
- Out of which circuit switching is the function of physical layer.

Medium :

- The signal always travels over some medium from sender to destination.
- The medium can be a coaxial cable or optical fiber etc. A medium does not pass all frequencies equally due to its inadequate frequency spectrum.
- It may pass some frequencies and weaken or block the other frequencies.
- Hence when a composite signal is passed over such a transmission medium, at the receiving end we get a wave, having a different shape as shown in Fig. 2.4.2.

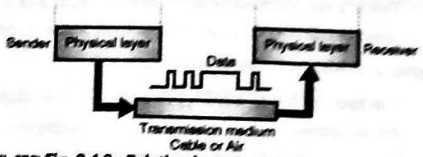


(G-92) Fig. 2.4.2 : Signal distortion on a transmission medium

- To avoid the signal distortion, the medium should pass all the frequencies present at the input without any change.
- But no medium is perfect and so some signal distortion is bound to take place.

2.4.2 Transmission Media and Physical Layer :

- Fig. 2.4.3 shows the location of the transmission media in the OSI model.



(G-570) Fig. 2.4.3 : Relation between transmission medium and physical layer

- So we can say that transmission media belongs to the lowest layer (layer 0) of the OSI model, as shown in Fig. 2.4.3 and the physical layer controls the transmission media directly.
- Transmission medium is the cable or air over which data can travel from sender to receiver in the form of digital or analog signals.

2.5 Transmission Media :

Definition :

- A transmission / communication media is defined as the medium over which information travels from the sender to receiver. A communication channel is also called as a medium.
- Different media have different properties and used in different environments for different purposes.
- The purpose of the physical layer is to transport a raw bit stream from one computer to another.

2.5.1 Classification of Transmission Media :

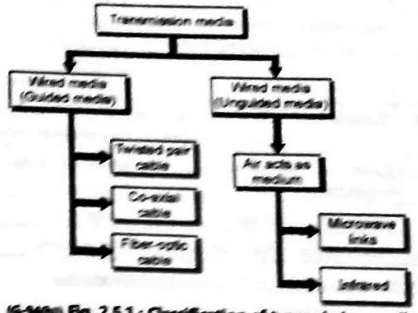
MU - May 17, May 19

University Questions

Q. 1 What are the different guided and unguided transmission media ? (May 17, 5 Marks)

Q. 2 Classify transmission media and compare them. (May 19, 10 Marks)

- The classification of transmissionmedia is as shown in Fig. 2.5.1.



(G-548p) Fig. 2.5.1 : Classification of transmission media

- Communication media are roughly grouped into two classes :

1. Wired media 2. Wireless media

Wired (guided) media :

- In this type of media, the signal energy is contained and guided within a solid media.
- The examples of wired media are copper pair wires, coaxial cables and fiber optic cables.
- The wired media is used for point to point communication.

Wireless (unguided) media :

- In the wireless media, the signal energy propagates in the form of unguided electromagnetic waves.
- The examples of wireless media are radio waves and infrared light.
- The wireless media is used for radio broadcasting in all the directions.

2.5.2 Comparison of Wired and Wireless Media :

MU - May 19

University Questions

Q. 1 Classify transmission media and compare them. (May 19, 10 Marks)

Sr. No.	Parameter of comparison	Wired media	Wireless media
1.	Principle	The signal energy is contained & guided within a solid medium.	The signal energy propagates as unguided E.M. waves.

Sr. No.	Parameter of comparison	Wired media	Wireless media
2.	Examples	Twisted pair cable, co-axial cable & fiber optic cable	Radio waves, microwaves, IR light.
3.	Applications	Point to point communication	Radio, TV broadcasting, satellite communication.
4.	Suitable for	Short distance communication	Long distance communication.
5.	Ease of installation	Not easy	Relatively easy.
6.	Bandwidth	Small	Large
7.	Maintenance	Frequent	Not required

2.5.3 Selection of Transmission Media :

- The important factors to be considered while selecting the transmission media are as follows :
  1. Type of medium (wired or wireless).
  2. Number of conductors.
  3. Flexibility.
  4. Durability or life span.
  5. Reliability of connection.
  6. Bandwidth.
  7. Effect of external interference.

2.6 Types of Wired Media : MU - Dec. 19

University Questions

Q. 1 Discuss different types of guided media in detail. **(Dec. 19, 10 Marks)**

- The most commonly used wired media are :
  1. Co-axial cable
  2. Twisted pair cable
  3. Optical fiber cable.
- The selection of networking media depends on various factors such as cost, connectivity, bandwidth, performance in presence of noise, geographical coverage etc.

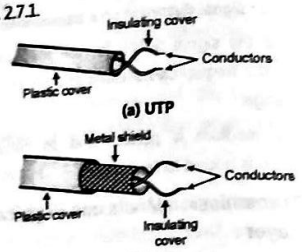
2.7 Twisted Pair Cables : MU - Dec. 19

University Questions

Q. 1 Discuss different types of guided media in detail. **(Dec. 19, 10 Marks)**

Construction :

- The construction of twisted pair cables is as shown in Fig. 2.7.1.



(L-574(A)) **Fig. 2.7.1 : Construction of twisted pair cables**

- This is a very commonly used wired medium and it is cheaper than the co-axial cable or optical fiber cable.

Types of twisted pair cables :

- The two commonly used types of twisted pair cables are as follows :
  1. Unshielded twisted pair (UTP)
  2. Shielded twisted pair (STP)
- The construction of UTP and STP cables is shown in Fig. 2.7.1.

2.7.1 UTP (Unshielded Twisted Pair) :

Construction :

- A twisted pair consists of two insulated conductors twisted together in the shape of a spiral as shown in Fig. 2.7.1.
- It can be either shielded or unshielded.
- The unshielded twisted pair (UTP) cables are very cheap and easy to install.
- But they are badly affected by the electromagnetic noise interference.

Characteristics of UTP :

1. The insulates wires are no shielded, but they are twisted around each other.
2. Noise and electromagnetic interference is high.
3. UTP is an economical guided medium.
4. Installation is easy

Transmission characteristics :

1. Noise and electromagnetic interference is high.
2. It has a low to moderate bandwidth.
3. It can be used only for point to point communication.
4. Supports data rates upto several Mbps.

Disadvantages :

1. Moderate data rates.
2. Low or moderate bandwidth.
3. Noise and EM interference is high.

Applications of UTP cables :

- Some of the applications of twisted pair cables are as follows :
  1. Local area networks for connecting computers to each other.
  2. In the analog telephony (conventional telephone line) to carry voice and data signals.
  3. In digital telephony system.

2.7.2 STP (Shielded Twisted Pair) :

Construction :

- A Shielded Twisted Pair (STP) cable is as shown in Fig. 2.7.1(b).
- It has a metal foil or braided mesh included in order to cover each pair of twisted insulating conductors.
- This is known as the metal shield which is normally connected to ground so as to reduce the interference of the noise. But this makes the cable bulky and expensive.
- Therefore practically UTP is more used than STP. The STP was developed by IBM and is used primarily for the IBM company only.
- Twisted pair cables can be used for either analog or digital transmission.
- The bandwidth supported by it depends on the thickness of the wire and the distance to be travelled by a signal on it.
- Twisted pairs support several megabits/sec for a few kilometers and are less costly.

Why to twist the wires ?

- Twisting of wires will reduce the effect of noise or external interference.
- The induced emf into the two wires due to interference tends to cancel each other if they are twisted.
- Number of twists per unit length will determine the quality of cable. More twists means better quality.

Physical characteristics :

1. STP uses two insulated conducting wires twisted around each other.
2. These twisted conductors are shielded by a braided mesh.
3. It is a low cost guided medium.

Transmission characteristics :

1. The noise and electromagnetic interference is low due to shielding and twisting of wires.
2. It supports data rates upto several Mbps.
3. It can be used only for point to point communication.
4. It has a low to moderate bandwidth.

Features of STP :

1. Low noise / electromagnetic interference.
2. Low cost medium.
3. Supports data rates up to few Mbps.
4. Low or moderate bandwidth.

Applications :

1. STP is used only for IBM computers.
2. Applications of the twisted pair cables are in point to point and point to multipoint communications, telephone systems etc.

2.7.3 Comparison of Twisted Pair Cables :

Sr. No.	Parameters	UTP	STP
1.	Bandwidth	1 – 155 Mbps (typically 10 Mbps)	1 – 155 Mbps (typically 16 Mbps)
2.	Number of node connected per segment	2	2
3.	Attenuation	High	High
4.	Electromagnetic interference	Very high	Low
5.	Easy of installation	Easy	Fairly easy
6.	Cost	Lowest	Moderate
7.	Speed	Lower than STP	Higher than UTP
8.	Security	Low	Moderate

2.8 Co-axial Cables :

University Questions

Q. 1 State different physical media properties, explain any one in detail. (May 07, 10 Marks)

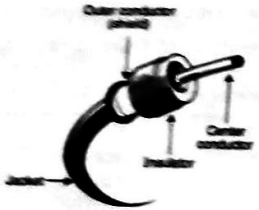
Q. 2 Discuss different types of guided media in detail. (Dec. 19, 10 Marks)

Construction :

- The construction of co-axial cable is as shown in Fig. 2.8.1.

(6-108) **Fig. 2.8.1 : Construction of a co-axial cable**

- It consists of two concentric conductors namely an inner conductor and a braided outer conductor separated by a dielectric material.
- The external conductor is in the form of metallic braid and used for the purpose of shielding.
- The co-axial cable may contain one or more co-axial pairs.
- The construction of a co-axial cable with other accessories such as connector, jacket etc. is shown in Fig. 2.8.2.



(6-579) **Fig. 2.8.2 : Co-axial cable**

- The wire mesh (braided conductor) protects the inner conductor from Electromagnetic Interference (EMI). It is often called a shield.
- A tough plastic jacket forms the cover of the cable as shown in Fig. 2.8.2 providing insulation and protection.
- The co-axial cable was initially developed for analog telephone networks. A single co-axial cable would be used to carry more than 10,000 voice channels at a time.
- The digital transmission systems using the co-axial cable were developed in 1970s.

Characteristics of a co-axial cable :

- The important characteristics of a co-axial cable are as follows :
  1. Due to the shield provided, this cable has excellent noise immunity.
  2. It has a large bandwidth and low losses.
  3. This cable is suitable for point to point or point to multipoint applications.
  4. These cables are costlier than twisted pair cables but they are cheaper than the optical fiber cables.
  5. The signal attenuation is less as compared to the twisted pair cable.
  6. Co-axial cables are easy to install.
  7. Co-axial cables are relatively inexpensive (as compared to the optical fiber cable) but they are costlier than twisted pair cables.

Advantages of co-axial cable :

1. Excellent noise immunity due to the shield.
2. Large bandwidth.
3. Losses are small.
4. Can support high data rates.
5. Less attenuation.
6. Ease of installation.

Disadvantages :

1. Costlier than the twisted pair cables.
2. BNC connectors are required to be used for connection.

Applications of co-axial cables :

1. Analog telephone networks.
2. Digital telephone network.
3. Cable TV.
4. Traditional Ethernet LANs.
5. Digital transmission.
6. Fast Ethernet.

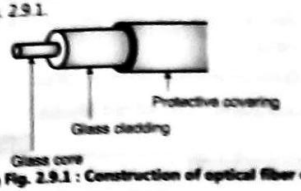
2.9 Optical Fiber Cables : MU : Dec. 19

University Questions

Q. 1 Discuss different types of guided media in detail. (Dec. 19, 10 Marks)

Construction :

- The construction of an optical fiber cable is as shown in Fig. 2.9.1.

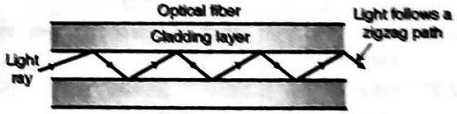


(6-103) **Fig. 2.9.1 : Construction of optical fiber cable**

- It consists of an inner glass core surrounded by a glass cladding which has a lower refractive index and a protective covering.
- Digital signals are transmitted in the form of intensity - modulated light signal.
- Light is launched into the fiber at one end using a light source such as a light emitting diode (LED) or laser.
- It is detected on the other side using a photo detector such as a phototransistor or photodiode.
- The optical fiber cables are costlier than the other two types but they have many advantages over the other two types.

Principle of light propagation in a fiber :

- The light enters into a glass fiber from one end, and gets reflected within the fiber.
- It follows a zigzag path along the length of the fiber as shown in Fig. 2.9.2.



(6-105) **Fig. 2.9.2. : Light follows a zigzag path within the optical fiber**

- Fig. 2.9.2 illustrates the principle of light travel through the optical fiber.
- When the light enters into a glass fiber from one end, most of it propagates along the length of the fiber and comes out from the far end.
- A small portion of the incident light escapes through the side walls of the fiber.
- The light which travels from one end to the other end of the glass fiber is said to have "guided" through the fiber.
- The light stays inside the fiber and does not escape through the walls because of the **"total internal reflection"** taking place inside the fiber.
- This total internal reflection (TIR) can take place only if the following two conditions are satisfied :

Conditions for TIR :

1. The glass fiber core must have a refractive index which is higher than the refractive index of the cladding around the core ($n_1 > n_2$).
2. The angle of incidence of the light entering the fiber must be greater than the **critical angle** "ϕ_c".

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

2.9.1 Characteristics of Optical Fiber Cables :

Physical characteristics :

1. Optical fibers are made from glass.
2. The information in the form of light travels over the optical fiber cables.
3. They are guided type media.
4. They are much expensive than the cables.
5. Installation is not easy.

Transmission characteristics :

1. Extremely large bandwidth (upto 2 Gbps).
2. High speed.
3. No effect of electromagnetic interference.
4. The number of nodes connected to this cable does not depend on the length.
5. They offer much lower attenuation. Hence a signal can travel very long distance without the need of any repeater or amplifier.
6. Three wavelength bands are used for fiber optic communication respectively 850 nanometer, 1300 nanometer, 1550 nanometer.
7. Fiber optic cable supports 75 nodes in an Ethernet network.
8. Single mode fiber optic cable are used to provide network links of several hundred kilometres in length.
9. Fiber optic cable does not leak signals so it is immune to evesdropping (tapping of signals).
10. Fiber optic cable does not require a ground, hence it is not affected by potential shifts in the electrical ground, nor does it produce sparks.

2.9.2 Advantages of Optical Fibers :

MU : Dec. 16, Dec. 18

University Questions

Q. 1 List the advantages of fiber optics as a communication medium. **(Dec. 16, Dec. 18, 5 Marks)**

- Some of the advantages of fiber optic communication over the conventional means of communication are as follows :

1. **Small size and light weight :**
- The size (diameter) of the optical fibers is very small (it is comparable to the diameter of human hair).
- Therefore a large number of optical fibers can fit into a cable of small diameter.
2. **No electrical or electromagnetic interference :**
- Since the transmission takes place in the form of light rays the signal is not affected due to any electrical or electromagnetic interference.
3. **Large bandwidth :**
- As the light rays have a very high frequency in the GHz range, the bandwidth of the optical fiber is extremely large.
- This allows transmission of more number of channels.
- Therefore the information carrying capacity of an optical fiber is much higher than that of a co-axial cable.
4. **Other advantages :**
- In addition to the advantages discussed earlier, the optical fiber communication has the following other advantages :
- No cross-talk inside the optical fiber cable.
- Signals at higher data rates can be sent.
- Intermediate amplifier are not required as the transmission losses in the fiber are low.
- These cables are not affected by the drastic environmental conditions.

2.9.3 Disadvantages of Optical Fiber :

- Some of the disadvantages of optical fibers are :

1. Sophisticated plants are required for manufacturing optical fibers.
2. The initial cost incurred is high.
3. Joining the optical fibers is a difficult job.

2.9.4 Applications :

1. Optical fibers are now used in the telephone systems.
2. In the Local Area Networks (LANs).

2.9.5 Comparison of Wired Media :

MU : Dec.

University Questions

Q. 1 Compare the performance characteristics of coaxial, twisted pair and fiber optic transmission media. **(Dec. 06, 10 Marks)**

Sr. No.	Parameter of comparison	Twisted pair cable	Co-axial cable	Fiber optic cable
1.	Transmitted Signal	Electrical	Electrical	Optical
2.	Noise / immunity	Low	High	Very high
3.	Effect of external magnetic field	Yes	Less	No effect
4.	Conductor short circuit	Possible	Possible	Not applicable
5.	Band width	Low	Moderate	Very large
6.	Attenuation	High	Medium	Low
7.	Ease of installation	Easy	Easy	Difficult
8.	Cost	Low	Medium	High

Review Questions

- **Q. 1** Name the layer which is associated with the transmission media.
- **Q. 2** What is the difference between guided and unguided transmission media ?
- **Q. 3** State the types of guided media.
- **Q. 4** Explain the difference between UTP and STP.
- **Q. 5** What is the effect of twisting the wires in UTP cables ?
- **Q. 6** Give applications of co-axial cable.
- **Q. 7** What is the advantage of using shielding ?
- **Q. 8** Compare the guided transmission media.
- **Q. 9** State advantages of optical fiber cable.
- **Q. 10** Compare twisted pair (UTP and STP).
- **Q. 11** Compare twisted pair, co-axial and fiber optic cable.

Module 3

Chapter 3

Data Link Layer

Syllabus

DLL design issues (Services, Framing, Error control, Flow control), Error detection and correction (Hamming code, CRC, Checksum), Elementary data link layer protocols, Stop and wait, Sliding window (Go back-N, Selective repeat).

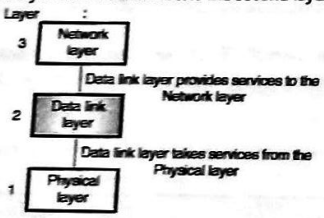
Chapter Contents

3.1 Introduction :

- The physical layer deals with the transmission of signals over different transmission media.
- A reliable and efficient communication between two adjacent machines can be achieved via the data link layer.
- This layer basically deals with frame formation, flow control, error control, addressing and link management.
- While sending data from source to destination errors may get introduced.
- The data communication circuits have only a finite data rate and there is non-zero propagation delay between the instant a bit is sent and the instant at which it is received.
- These limitations affect the efficiency of data transfer.
- The data link layer protocols used for communication take care of all these problems.
- Data link layer is the second layer in OSI reference model. It is above the physical layer.

3.1.1 Position of Data Link Layer :

- Fig. 3.1.1 shows the position of data link layer in the five layer Internet model. It is the second layer.



(L-663) Fig. 3.1.1 : Position of data link layer

- It receives services from the physical layer and provides services to the network layer.

3.2 Data Link Layer Design Issues:

MU - May 10, May 11, May 16, May 17, May 19

University Questions

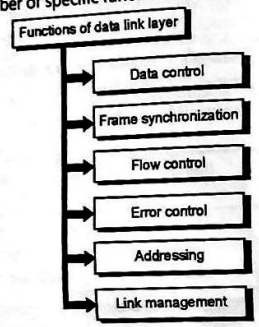
Q. 1 Describe any five functions of data link layer with suitable examples. (May 10, May 11, 10 Marks)

Q. 2 Explain any four functions of data link layer with example. (May 16, 10 Marks)

Q. 3 Enumerate the main responsibilities of the data link layer. (May 17, 5 Marks)

Q. 4 Explain design issues of Data Link layer. Explain Sliding Window protocol Selective Repeat. (May 19, 10 Marks)

- The data link layer is supposed to carry out many specified functions.
- For effective data communication between two directly (physically) connected transmitting and receiving stations the data link layer has to carry out a number of specific functions as follows :



(L-664)Fig. 3.2.1 : Functions of data link layer

1. **Services provided to the network layer :**
- *The data link layer provides a well defined service interface to the network layer.*
- *The principle service is transferring data from the network layer on sending machine to the network layer on destination machine.*
- *This transfer always takes place via the DLL.*

2. **Frame synchronization :**
- The source machine sends data in the form of blocks called frames to the destination machine.
- The starting and ending of each frame should be identified so that the frames can be recognized by the destination machine.

3. **Flow control :**
- The source machine must not send data frames at a rate faster than the capacity of destination machine to accept them.

4. **Error control :**
- The errors introduced during transmission from source to destination machines must be detected and corrected at the destination machine.

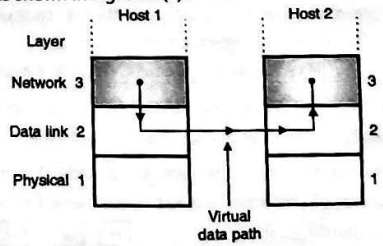
5. **Addressing :**
- When many machines are connected together (LAN), the identity of the individual machines must be specified while transmitting the data frames. This is known as addressing.

6. **Control and data on same link :**
- The data and control information is combined in a frame and transmitted from the source to destination machine.
- The destination machine must be able to separate out the control information from the data being transmitted.

7. **Link management :**
- The communication link between the source and destination is required to be initiated, maintained and finally terminated for effective exchange of data.
- It requires co-ordination and co-operation among all the involved stations.
- Protocols or procedures are required to be designed for the link management.

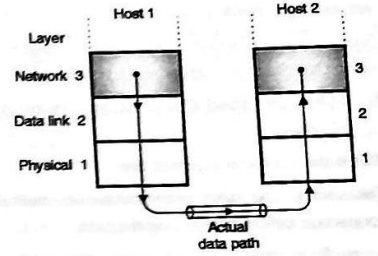
3.3 Services Provided to Network Layer :

- Network layer is the layer above the data link layer in the OSI model.
- So it is supposed to provide services to the network layer.
- The main service to be provided is to transfer data from the network layer on the sending machine to the network layer of the receiving machine.
- The virtual path followed for such a communication is shown in Fig. 3.3.1(a).



(L-665) Fig. 3.3.1(a) : Virtual communication

- It is not the actual path.
- The actual path followed by the data from sending machine to destination is shown in Fig. 3.3.1(b) which is via all the layers below the network layer, then the physical medium, then layers 1, 2, 3 of receiving machine.



(L-665) Fig. 3.3.1(b) : Actual data path

- However it is always easier to think that the communication is taking place through the data link layers (Fig. 3.3.1(a)) using a data link layer protocol.

3.3.1 Types of Services Provided :

- Data link layer can be designed to offer different types of services. Some of them are as follows :
  1. Unacknowledged connectionless service.
  2. Acknowledged connectionless service.
  3. Acknowledged connection oriented service.

3.3.2 Unacknowledged Connectionless Service :

- In this type of service, the destination machine does not send back any acknowledgement after receiving frames.
- It is a connectionless service. So no connection is established before communication or released after it is over.
- If a frame is lost due to channel noise, then there are no attempts made to recover it.
- So this service is suitable only if the error rate is low.
- It is suitable for real time traffic such as speech. This type of service is highly unreliable.

3.3.3 Acknowledged Connectionless Service :

- This is the next step to improve reliability.

- In this service, there are no connections established for data transfer but for each frame received, the receiver sends an acknowledgement to the sender.
- If a frame is not received within some specified time it is assumed to be lost and the sender will retransmit it.
- This service is suitable for communication over unreliable channels such as wireless channels.

3.3.4 Acknowledged Connection Oriented Service :

- This is the most sophisticated one.
- The source and destination machines establish a connection before transferring the data.
- A specific number is given to each frame being sent and the data link layer guarantees that each transmitted frame is received.
- All the frames are guaranteed to be received in the same order as the order of transmission.
- Each received frame will be acknowledged individually by the destination machine.

Three phases :

- The data transfer takes place by following three distinct phases given below :
  1. Connection is established.
  2. The data frames are actually transmitted.
  3. The connection is released after completion of data transfer.

3.4 Framing :

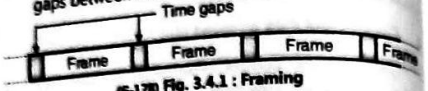
MU : Dec. 08, Dec. 09, May 10

University Questions

Q. 1 Explain framing, flow and error control in data link layer. **(Dec. 08, Dec. 11, 10 Marks)**

Q. 2 Explain any four functions of data link layer with example. **(May 16, 10 Marks)**

- The bits to be transmitted are first broken into discrete frames at the data link layer.
- In order to guarantee that the bit stream is error free, the checksum of each frame is computed.
- When a frame is received, the data link there, recomputes the checksum.
- If it is different from the checksum present in the frame, then the data link layer knows that an error has occurred.
- It then discards the bad frame and sends back a request for retransmission.
- Breaking the bit stream into frames is called as framing. One way of doing it is by inserting time gaps between frames as shown in Fig. 3.4.1.



(6-178) **Fig. 3.4.1 : Framing**

- However, practically this framing technique does not work satisfactorily, because networks generally do not make any guarantees about the timing.
- Therefore, some other methods are derived.

3.4.1 Framing Methods :

MU : May 13, May 15, May 16, Dec. 17, Dec. 18, May 19

University Questions

Q. 1 Explain different framing methods. What are the advantages at variable length frames over fixed length frames ? **(May 13, May 19, 10 Marks)**

Q. 2 Explain in short different framing methods. **(May 15, May 16, Dec. 17, Dec. 18, 4 Marks)**

- Following methods are used for carrying out framing :
  1. Character count method.
  2. Starting and ending characters, with character stuffing.
  3. Starting and ending flags with bit stuffing.
  4. Physical layer coding violations.

3.4.2 Character Count :

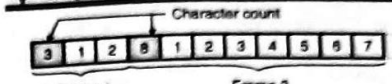
MU : May 13, May 15, May 16, Dec. 17, Dec. 18, May 19

University Questions

Q. 1 Explain different framing methods. What are the advantages at variable length frames over fixed length frames ? **(May 13, May 19, 10 Marks)**

Q. 2 Explain in short different framing methods. **(May 15, May 16, Dec. 17, Dec. 18, 4 Marks)**

- In this method, a field in the header is used to specify the number of characters in the frame.
- This number helps the receiver to know the exact number of characters present in the frame following this count.
- The character count method is illustrated in Fig. 3.4.2.



(L-668) **Fig. 3.4.2 : Character count method**

- The two frames shown in Fig. 3.4.2 contain 3 and 8 characters respectively and numbers 3 and 8 are inserted in the headers of the corresponding frames.
- The disadvantage of this method is that, an error can change the character count itself.
- If the wrong character count number is received due to error then the receiver will get out of synchronization and will not be able to locate the start of next frame.
- The character count method is rarely used in practice.

3.4.3 Starting and Ending Character with Character Stuffing :

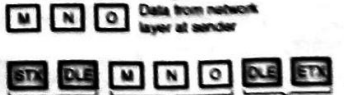
MU : May 13, May 15, May 16, Dec. 17, Dec. 18, May 19

University Questions

Q. 1 Explain different framing methods. What are the advantages at variable length frames over fixed length frames ? **(May 13, May 19, 10 Marks)**

Q. 2 Explain in short different framing methods. **(May 15, May 16, Dec. 17, Dec. 18, 4 Marks)**

- The problem of character count method is solved here by using a starting character before the starting of each frame and an ending character at the end of each frame.
- Each frame is preceded by the transmission of ASCII character sequence DLE STX. (DLE stands for data link escape and STX is start of Text).
- After each frame the ASCII character sequence DLE ETX is transmitted.
- Here DLE stands for Data Link Escape and ETX stands for End of Text. So if the receiver loses t
- he synchronization, it just has to search for the DLE STX or DLE ETX characters to return back on track. This is shown in Fig. 3.4.3.



(L-669)**Fig. 3.4.3**

3.4.4 Character Stuffing :

MU : May 11, May 13, May 15, Dec. 17, Dec. 18, May 19

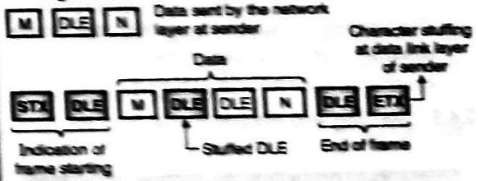
University Questions

Q. 1 Explain the different framing methods. **(May 11, Dec. 17, 5 Marks)**

Q. 2 Explain different framing methods. What are the advantages at variable length frames over fixed length frames ? **(May 13, May 19, 10 Marks)**

Q. 3 Explain in short different framing methods. **(May 15, May 16, Dec. 17, Dec. 18, 4 Marks)**

- The problem with this system is that the characters DLE STX or DLE ETX can be a part of data as well.
- If so, they will be misinterpreted by the receiver as start or end of frame.
- This problem is solved by using a technique called **character stuffing** which is as follows.
- The data link layer at the sending end inserts an ASCII DLE character just before each accidental DLE character in the data being transmitted.
- The data link layer at the receiving end will remove these DLE characters before transferring the data to the network layer.
- Thus the DLE STX or DLE ETX used for framing purpose can be distinguished from the one in data because DLEs in the data always appear more than once.
- This is called character stuffing and it is shown in Fig. 3.4.4.



(6-182) **Fig. 3.4.4 : Character stuffing**

- Note that at the receiving end the destuffing is essential. Destuffing process is exactly opposite to the character stuffing process.

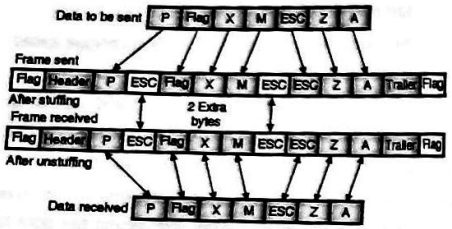
Disadvantages :

- The main disadvantage of this framing method is that we have to use the 8 bit characters and ASCII code.

- This problem can be overcome by using the next framing technique.

Byte stuffing :

- In byte stuffing a special byte is added to the data section of the frame when there is a character with the same pattern as the flag.
- The data section is stuffed with an extra byte. This byte is called as the escape character (ESC).
- At the receiver these ESC bytes are removed from the data section and the next character is treated as data.
- Fig. 3.4.5 demonstrates the concept of byte stuffing.



(G-182) **Fig. 3.4.5 : Byte-stuffing**

- Byte stuffing by the escape character will allow the presence of the flag is the data section of the frame.
- But it has a problem, if the text contains one or more escape characters followed by a flag.
- Because then the receiver will remove the escape character but will keep the flag.
- This problem is solved by marking the escape characters that are a part of the text by another escape (ESC) character as shown in Fig. 3.4.5.

3.4.5 Starting and Ending Flags, with Bit Stuffing :

MU : May 13, May 15, May 16, Dec. 17, Dec. 18, May 19

University Questions

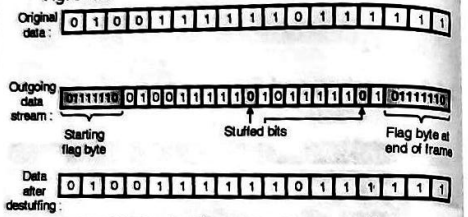
Q. 1 Explain different framing methods. What are the advantages at variable length frames over fixed length frames ? **(May 13, May 19, 10 Marks)**

Q. 2 Explain in short different framing methods. **(May 15, May 16, Dec. 17, Dec. 18, 4 Marks)**

- In this framing techniques at the beginning and end of each frame, a specific bit pattern 0111 1110 called **flag byte** is transmitted by the sending station.
- Since there are six consecutive 1s in the flag byte a technique called **bit stuffing** which is similar to character stuffing is used. It is as explained below.

Bit stuffing :

- Whenever the sender data link layer detects the presence of five consecutive ones in the data stream, it automatically stuffs a 0 bit into the outgoing bit stream.
- Thus the six consecutive 1s will never appear in the data stream. Hence there is no chance of misinterpretation.
- This is called bit stuffing and it is illustrated in Fig. 3.4.6.



(G-183) **Fig. 3.4.6 : Bit stuffing and destuffing**

- When a receiver detects presence of five consecutive ones in the received bit stream, it automatically deletes the 0 bit following the five ones.
- This is called de-stuffing. It is shown in Fig. 3.4.6.
- Due to bit stuffing, the possible problem if the data contains the flag byte pattern (0111 1110) is eliminated.

3.4.6 Physical Layer Coding Violations :

MU : Dec. 10, May 11, May 13, May 16, Dec. 17

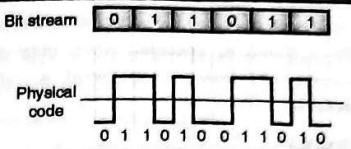
University Questions

Q. 1 Explain the different framing methods. **(Dec. 10, May 11, 5 Marks)**

Q. 2 Explain different framing methods. What are the advantages at variable length frames over fixed length frames ? **(May 13, 10 Marks)**

Q. 3 Explain in short different framing methods. **(May 16, Dec. 17, 4 Marks)**

- This method of framing is applicable only to those networks in which the encoding on the physical medium contains some redundancy.
- Some LANs encode each bit of data using two physical bits for example the use of the Manchester coding refer Fig. 3.4.7.



(G-184) **Fig. 3.4.7**

- The physical Manchester code makes a transition at the middle of the bit interval as shown.
- Therefore a 1 bit is encoded into a 10 pair and a 0 bit is encoded into a 01 pair as shown in Fig. 3.4.7. This helps in recognizing the boundaries of bits in a precise manner.
- This use of invalid physical code is a part of 802 LAN standards.

Which method of framing is used practically ?

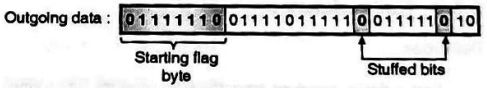
- Many data link protocols use the combination of the character count technique with one of the other techniques so as to have an extra safety.

Ex. 3.4.1 : A bit string 0111101111101111110, needs to be transmitted at the data link layer. What is the string actually transmitted after bit stuffing ?

Soln. :

- The original bit stream and the stream after bit stuffing are shown in Fig. P. 3.4.1.

Original data : 01111011111011111110



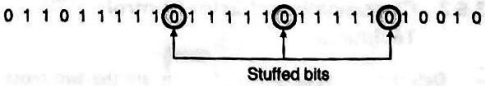
(G-217) **Fig. P. 3.4.1**

Ex. 3.4.2 : Apply bit stuffing

01101111111111111111110010

Soln. :

- The outgoing data after bit stuffing is shown in Fig. P. 3.4.2.



(G-218) **Fig. P. 3.4.2**

3.5 Error Control :

MU : Dec. 09, Dec. 11

University Questions

Q. 1 Explain framing, flow and error control in data link layer. **(Dec. 09, Dec. 11, 10 Marks)**

- The next problem to be dealt with is to make sure that all frames are eventually delivered to the network layer at the destination, in proper order.
- Generally the receiver sends back some feedback (positive or negative) to convey the information about whether it has received a frame or not.
- A positive acknowledgement (feedback) ACK indicates a successful and error free delivery of a frame.
- Whereas a negative acknowledgement (NAK) means that something has gone wrong and that particular frame needs to be retransmitted.
- Due to the presence of noise burst a frame may vanish completely.
- So the receiver does not receive anything and it does not react at all (no acknowledgement).
- This problem is overcome by introducing a timer in the data link layer. Its function of this timer is as follows.

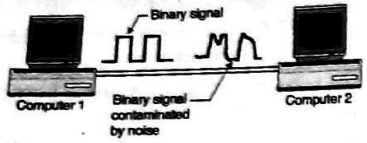
3.5.1 Function of a Timer :

- As soon as a sender transmits a frame, it also starts the data link timer.
- The timer timing is set by taking into account the factors such as the time required for the frame to reach the destination, processing time at the destination and the time required for the acknowledgement to return back.
- Normally the frame is received correctly and the acknowledgement will return back to the sender before the timer runs out.
- This shows that a frame has been received and the timer is cancelled.
- But if a frame is lost or acknowledgement is lost, then the timer will go off. This will alert the sender that there is some problem.
- The solution to this problem is that the sender retransmits the same frame.
- But when a frame is transmitted multiple times, there is a possibility that the receiver will receive the same frame two or more times and pass it to the network layer more than once. This is called as duplication.

- To avoid this each outgoing frame is assigned a distinct sequence number.
- This will help the receiver to distinguish retransmission.

3.6 Error Detection and Correction :

- When transmission of digital signals takes place between two systems such as computers as shown in Fig. 3.6.1, the signal get contaminated due to the addition of "Noise" to it.

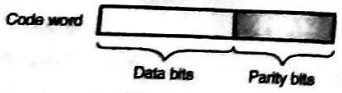


(L-302) **Fig. 3.6.1 : Noise contaminates the binary signal**

- The noise can introduce an error in the binary bits travelling from one system to the other.
- That means a 0 may change to 1 or a 1 may change to 0.
- These error can become a serious threat to the accuracy of the digital system.
- Therefore it is necessary to detect and correct the errors.

3.6.1 Encoding and Decoding :

- For the detection, and / or correction of errors, one or more than one, extra bits are added to the data bits at the time of transmitting.
- These extra bits are called as **parity** bits or **redundant** bits.
- They do not contain any information. They are produced from the data bits using some predefined rules.
- However, they allow the detection or sometimes correction of the errors.
- The data bits along with the parity bits form a **code word** as shown in Fig. 3.6.2.

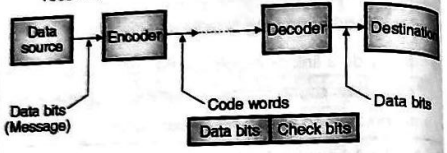


(L-303) **Fig. 3.6.2 : Structure of a transmitted code word**

- The process of adding the parity bits to the message bits is known as **encoding** while, obtaining the message bits from a code word is called as **decoding**.

Block diagram :

- Fig. 3.6.3 shows the encoding and decoding processes.
- The source generates the data (message) in the form of binary symbols.
- The encoder accepts these bits and adds the check (parity) bits to them to produce the code words.
- These code words are transmitted towards the receiver.



(L-306) **Fig. 3.6.3 : Encoding and decoding**

- The decoder separates out the data and check bits. It uses the parity bits to detect and correct errors if they are present in the received code words.
- The data bits are then passed on to the destination.

3.6.2 Redundancy :

Definition :

- Redundancy involves transmission of extra bits along with the data bits.
- These extra bits actually do not contain any data or information but they ensure the detection and correction of errors introduced during the data travel from sender to receiver.

3.6.3 Classification of Error Control Techniques :

- Detection and correction of errors are the two most important aspects of error control in data communication.
- The error control techniques can be divided into two types as follows :
  1. Error detection techniques
  2. Error correction techniques

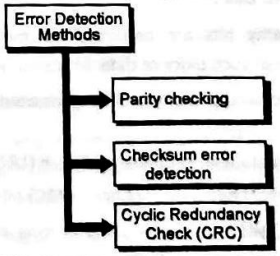
1. **Error detection :**
   - The error detecting techniques are capable of only detecting the errors.
   - They cannot correct the errors.
   - In **error detection** we are not interested even in the number of errors.
   - The only question to be answered is whether an error has occurred or not.
2. **Error correction :**
   - The error correcting techniques are capable of detecting as well as correcting the errors.
   - In **error correction**, multiple processes are involved such as detecting the errors, knowing their number, the location of errors and then correcting the erroneous bits.
   - The correction of errors is more difficult as compared to their detection.

3.7 Error Detection Techniques :

- A number of methods (techniques) are available for the detection and correction of errors introduced in the transmitted signal.
- When a codeword is transmitted, one or more number of transmitted bits will be reversed (0 to 1 or vice versa) due to transmission impairments. Thus errors are introduced.

Error detection techniques :

- Some of the most important error detection methods are as shown in Fig. 3.7.1.



(G-192) **Fig. 3.7.1: Classification of error detection techniques**

- For error detection and correction it is necessary to add some check bit to a block of data bits.
- These check bits are also known as redundant bits because they do not carry any useful information.

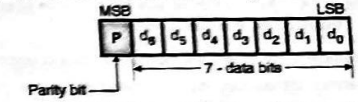
3.7.1 Parity Checking :

Definition of parity bit :

- A parity bit or a check bit is a bit added to a string of binary bits to ensure that the total number of 1-bit in the string including the parity bit is either even or odd.

Addition of parity bit :

- The simplest technique for detecting errors is to add an extra bit known as parity bit to each word being transmitted.
- As shown in Fig. 3.7.2, generally the MSB of an 8-bit word is used as the parity bit and the remaining 7 bits are used as data or message bits.

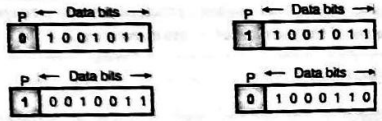


(L-309) **Fig. 3.7.2 : Format of a transmitted word with parity bit**

- The parity of the 8-bit transmitted word can be either even parity or odd parity. Even parity means the number of 1's in the given word including the parity bit should be even (2, 4, 6...).
- Odd parity means the number of 1's in the given word including the parity bit should be odd (1, 3, 5...).

Use of Parity Bit to Decide Parity :

- The parity bit can be set to 0 or 1 depending on the type of parity required.
- For odd parity this bit is set to 1 or 0 at the transmitter such that the number of "1 bits" in the entire word is odd.



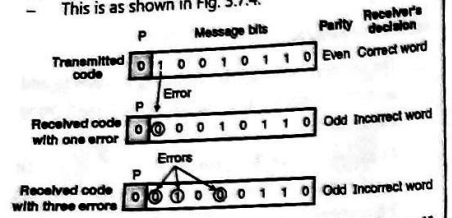
(a) Inclusion of a parity bit to obtain an even parity **(b) Inclusion of a parity bit to obtain the odd parity**

(L-310)**Fig. 3.7.3**

How does error detection take place ?

- The parity checking at the receiver can detect the presence of an error if the parity of the received signal is different from the expected parity.

- That means if it is known that the parity of the transmitted signal is always going to be "even" and if the received signal has an odd parity then the receiver can conclude that the received signal is not correct.
- This is as shown in Fig. 3.7.4.

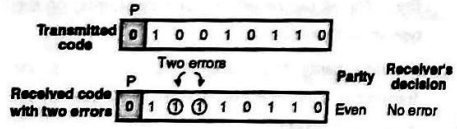


(L-311)Fig. 3.7.4 : The receiver detects the presence of error if the number of errors is odd i.e. 1, 3, 5 ....

- If a single error or an odd number of bits change due to errors introduced during transmission the parity of the code word will change.
- If presence of error is detected then the receiver will ignore the received byte and request for the retransmission of the same byte to the transmitter.

When does parity checking fail to detect errors ?

- If the number of errors introduced in the transmitted code is two or any even number, then the parity of the received code word will not change.
- It will still remain even as shown in Fig. 3.7.5 and the receiver will fail to detect the presence of errors.



(L-312)Fig. 3.7.5 : The receiver cannot detect the presence of error if the number of errors is even i.e. 2, 4, 6 ....

Conclusions :

1. *Double or any even number of errors in the received word will not change the parity. Therefore even number of errors will be unnoticed.*
2. *If one or odd number of errors occur then the parity of the received word will be different from the parity of transmitted signal. Thus error is noticed. However this error can neither be located nor be corrected.*

Limitations of parity checking :

- Following are the limitations of parity check technique :
  1. It is not suitable for detection of multiple errors (two, four, six etc).
  2. The other limitation of parity checking method is that it cannot reveal the location of erroneous bit.
  3. It cannot correct the error either.

3.7.2 Two Dimensional Parity Check (Block Parity) :

Block of Data :

- When a large number of binary words are being transmitted or received in succession, the resulting collection of bits is considered as a **block of data** with rows and columns as shown in Fig. 3.7.6.

Characters		C	O	M	P	U	T	E	R	
7 bit ASCII codes (Message bits)	b_1	1	1	1	0	1	0	1	0	1
	b_2	1	1	0	0	0	0	0	1	1
	b_3	0	1	1	0	1	1	1	0	1
	b_4	0	1	1	0	0	0	0	0	0
	b_5	0	0	0	1	1	1	0	1	0
	b_6	0	0	0	0	0	0	0	0	0
VRC bits (even parity)	b_7	1	1	1	1	1	1	1	1	0
		1	1	0	0	0	1	1	1	1

These bits will make the parity of each column even

These bits will make the parity of each row even ← LRC bits (even parity)

(L-315) Fig. 3.7.6 : Vertical and longitudinal parity check bits

LRC and VRC Bits :

- The parity bits are produced for each row and column of such block of data.
- The two sets of parity bits so generated are known as :
  1. Longitudinal Redundancy Check **(LRC)** bits
  2. Vertical Redundancy Check **(VRC)** bits.
- The LRC bits indicate the parity of rows and VRC bits indicate the parity of columns as shown in Fig. 3.7.6.

The Vertical Redundancy Check (VRC) Bits :

- As shown in Fig. 3.7.6 the VRC bits are parity bits associated with the ASCII code of each character.

- Each VRC bit will make the parity of its corresponding column "an even parity".
- For example consider column 1 corresponding to character "C".
- The ASCII code for the character C is as follows:

C = b_7..... b_1 = 1 0 0 0 0 1 1

Character	C
b_1	1
b_2	1 ← Column - 1 of the data block
b_3	0
b_4	0
b_5	0
b_6	0
b_7	1
VRC bit →	1 ← VRC bit = 1 to make the parity of first column even

(G-1944)

- Thus it has three 1's. Therefore, the 8th bit which is a VRC bit is made "1" in order to make the parity even.
- Similarly the other VRC bits are found as shown in Fig. 3.7.6.

The Longitudinal Redundancy Check (LRC) Bits :

- The LRC bits are parity bits associated with the rows of the data block of Fig. 3.7.6.
- Each LRC bit will make the parity of the corresponding row, an even parity.
- For example, consider row 1 of Fig. 3.7.6 as follows.

Row 1 : b_1 1 1 1 0 1 0 1 0 1 ← LRC bit to make parity even

(G-1945)

- The contents of first row are (1 1 1 0 1 0 1 0), which has five 1s.
- Hence, in order to make the parity of the first row even, we select the LRC bit = 1.

How to locate the bit in error ?

- Even a single error in any bit will result in an incorrect "LRC" in one of the rows and an incorrect VRC in one of the columns.
- The bit which is common to the row and column is the bit in error.
- However there is still a limitation on the block parity code, which is that, multiple errors in rows and columns can be only detected but they cannot be corrected.
- This is because, it is not possible to locate the bits which are in error.

3.7.3 Checksum Error Detection :

Definition :

- A checksum is a small-sized datum derived from a block of digital data for the purpose of detecting errors that may have been introduced during its transmission or storage.
- Checksum is the last error detection method. It is used in the Internet by many protocols.
- The concept of checksum is based on the principle of redundancy. This is very similar to linear block codes and cyclic codes.
- Some protocols still use checksum for error detection but in the recent days CRC is more preferred.
- Hence CRC is fast replacing the checksum.
- Checksum is used at the data link layer level and CRC is used in some layers other than the data link layer.
- Simple parity cannot detect two or even number of errors within the same word.
- One way to overcome this problem is to use a sort of two dimensional parity.

Calculation of checksum :

- As each word is transmitted, it is added to the previously sent word and the sum is retained at the transmitter as shown in Fig. 3.7.7.

Word A : 1 0 1 1 1 0 1 1 1
+
Word B : 0 0 1 0 0 0 1 0
Sum : 1 1 0 1 1 0 0 1

(L-898) Fig. 3.7.7 : Concept of checksum

- Each successive word is added in this manner to the previous sum.
- At the end of the transmission the sum (called a checksum) up to that time is sent.
- The errors normally occur in burst. The parity check method is not useful in detecting the errors under such conditions.
- The checksum error detection method can be used successfully in detecting such errors.
- In this method a "checksum" is transmitted along with every block of data bytes.

- In this method an eight bit accumulator is used to add 8 bit bytes of a block of data to find the "checksum byte".
- The carries of the MSB are ignored while finding out the checksum byte.
- The generation of checksum will be clear if you refer to the following example.

Ex. 3.7.1 : Find the checksum of the following message.

10110001, 10101011,

00110101, 10100001

Soln. :

Ignored

Carries	(1)0 1 0 1 1 1 1 0
Data bytes	+ 1 0 1 1 0 0 0 1 + 1 0 1 0 1 0 1 1 + 0 0 1 1 0 1 0 1 + 1 0 1 0 0 0 0 1
Checksum byte	0 0 1 1 0 0 1 0

(G-1943)

- Note that the carries of MSB have been ignored while writing the checksum byte.

How to detect error using the checksum byte ?

1. After transmitting a block of data bytes (say 8-data bytes) the "checksum" byte is also transmitted. The checksum byte is regenerated at the receiver separately by adding the received bytes.
2. The regenerated checksum byte is then compared with the transmitted one. If both are identical then there is no error. If they are different then the errors are present in the block of received data bytes.
3. Sometimes the 2's complement of the checksum is transmitted instead of the checksum itself. The receiver will accumulate all the bytes including the 2's complement of the checksum. If there is no error, the contents of the accumulator should be zero after accumulation of the 2's complement of the checksum byte.

Advantage of the Checksum Method :

- The advantage of this method over the simple parity checking method is that the data bits are "mixed up" due to the 8 bit addition.
- Therefore checksum represents the overall data block.
- In checksum therefore, there is 255 to 1 chance of detecting random errors.

Ex. 3.7.2 : What is the checksum of the following characters ?

01011010, 11000101,

1101 1001.

Soln. :

- The checksum is calculated as follows:

	0 1 0 1 1 0 1 0	← Byte 1
+	1 1 0 0 0 1 0 1	← Byte 2
+	1 1 0 1 1 0 0 1	← Byte 3
	(1) 1 1 1 1 1 1	← Byte 4
Discard final carry	1 1 1 1 1 0 0 0	← Checksum

(G-1986)

3.8 Cyclic Redundancy Check (CRC) :

MU : May 09, Dec. 12, May 13

University Questions

Q. 1 What is CRC ? Write the algorithm for computing checksum and explain with suitable example. **(May 09, 10 Marks)**

Q. 2 What is CRC ? **(Dec. 12, 5 Marks)**

Q. 3 Explain with the suitable example CRC algorithm for computing checksum. **(May 13, 10 Marks)**

Definition :

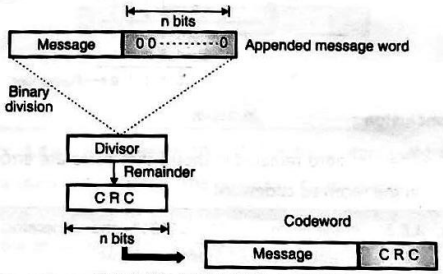
- *CRC is an error detection code which is included in each transmitted codeword as shown in Fig. 3.8.1 and used by the receiver to detect the errors in the received codeword.*
- *This is a type of polynomial code in which a bit string is represented in the form of polynomials with coefficients of 0 and 1 only.*
- *Polynomial arithmetic uses a modulo-2 arithmetic i.e. addition and subtraction are identical to EXOR.*
- *For CRC code the sender and receiver should agree upon a generator polynomial G(x).*
- *A codeword can be generated for a given data word (message) polynomial M(x) with the help of long division.*
- *This technique is more powerful than the parity check and checksum error detection.*

Procedure of error detection :

- CRC works on the principle of binary division. A sequence of redundant bits called CRC or CRC remainder is appended at the end of the message. We will call this word as appended message word.
- The appended word thus obtained becomes exactly divisible by the generator word corresponding to G (x).
- The sender appends the CRC to the message word to form a codeword.
- At the receiver, this codeword is divided by the same generator word which corresponds to G (x).
- There is no error if the remainder of this division is zero. But a non-zero remainder indicates presence of errors in the received codeword.
- Such an erroneous codeword is then rejected.

3.8.1 CRC generator :

- The CRC generator is shown in Fig. 3.8.1.



(L-819) **Fig. 3.8.1 : CRC generator**

- The stepwise procedure in CRC generation is as follows :

Step 1 : Append a train of n 0s to the message word where n is 1 less than the number of bits in the predecided divisor (i.e. generator word). If the divisor is 5-bit long then we have to append 4-zeros to the message.

Step 2 : Divide the newly generated data unit in step 1 by the divisor (generator). This is a binary division.

Step 3 : The remainder obtained after the division in step 2 is the n bit CRC.

Step 4 : This CRC will replace the n 0s appended to the data unit in step 1, to get the codeword to be transmitted as shown in Fig. 3.6.10.

Generation of CRC code :

- The generation of CRC code is clear after solving the following example.

Ex. 3.8.1 : Generate the CRC code for the data word of 1 1 0 0 1 0 1 0 1. The divisor is 1 0 1 0 1.

Soln. :

Given : Data word : 1 1 0 0 1 0 1 0 1

Divisor : 1 0 1 0 1.

The number of data bits = m = 9

The number of bits in the codeword = n = 5

Step 1 : Obtain the dividend :

Dividend = Data word + (n – 1) zeros.

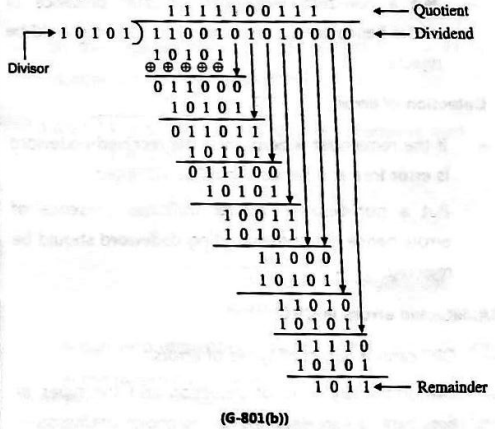
Dividend = 1 1 0 0 1 0 1 0 1 | 0 0 0 0

Data word — 4 additional zeros

(G-801(a))

Step 2 : Carry out the division :

- Carry out the division as follows :



(G-801(b))

Step 3 : Obtain the Codeword :

- In CRC the required codeword is obtained by writing the data word followed by the remainder.

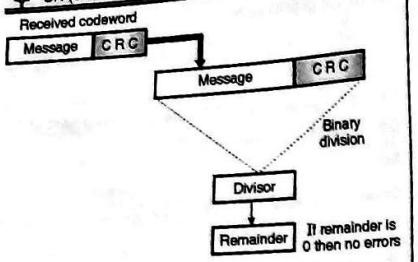
∴ Codeword = 1 1 0 0 1 0 1 0 1 | 1 0 1 1

Data word — Remainder

(G-1761)

3.8.2 CRC checker :

- Fig. 3.8.2 shows the CRC checker.



(L-820) **Fig. 3.8.2 : CRC checker**

- The codeword received at the receiver consists of message and CRC. (Fig. 3.8.2)
- The receiver treats it as one unit and divides it by the same (n + 1) bit divisor (generator word) which was used at the transmitter.
- The remainder of this division is then checked.
- If the remainder is zero, then the received codeword is error free and hence should be accepted.
- But a non-zero remainder indicates presence of errors hence the corresponding codeword should be rejected.

Detection of error :

- If the remainder is zero, then the received codeword is error free and hence should be accepted.
- But a non-zero remainder indicates presence of errors hence the corresponding codeword should be rejected.

Undetected errors in CRC :

- CRC cannot detect all types of errors.
- The probability of error detection and the types of detectable errors depends on the choice of divisor.

Ex. 3.8.2 : The codeword is received as 1100 1001 01011. Check whether there are errors in the received codeword, if the divisor is 10101. (The divisor corresponds to the generator polynomial).

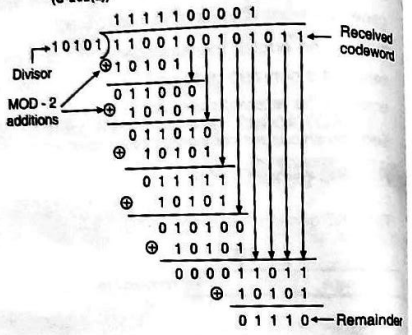
Soln. :

- As we know the codeword is formed by adding the dividend and the remainder.
- This codeword will have an important property that it will be completely divisible by the divisor.
- Thus at the receiver we have to divide the receiv codeword by, the same divisor and check for th remainder.
- If there is no remainder then there are no errors. Bu if there is remainder after division, then there ar errors in the received codeword.
- Let us use this technique and find if there are errors

Data word : 1100 1001 01011

Divisor : 10101

(G-201(a))



Conclusion :

- The non zero remainder shows that there are erro in the received codeword.

Ex. 3.8.3 : If the frame is 1101011011 and generator i $x^4 + x + 1$ what would be the transmitted frame.

May 05, Dec. 09, 5 Marks, May 1 Dec. 19, 10 Mark

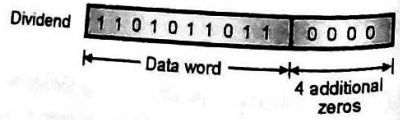
Soln. :

Given : Data word : 1101011011

Generator : $x^4 + x + 1 = x^4 + 0\,x^3 + 0\,x^2 + x + 1$ = 10011 = n

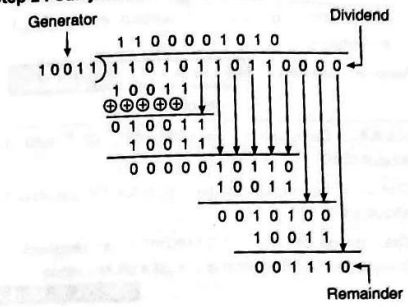
Step 1 : Add four zeros at the end of the data word :

- Add four zeros (n – 1) at the end of data word to ge the dividend as follows :



(G-216(c))

Step 2 : Carryout the long division :

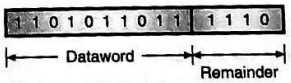


(G-216(d))

Step 3 : Write the transmitted frame :

- The transmitted frame is obtained by writing the data word followed by the remainder.

∴ The transmitted codeword is as follows :



(G-216(e))

Ex. 3.8.4 : A bit stream 10011101 is transmitted using the standard CRC method. The generator polynomial is $x^3 + 1$. Show the actual bit string transmitted. Suppose the third bit from left is inverted during transmission. Show that this error is detected at the receiver's end.

Soln. :

Given : Data word (Bit string) : 10011101

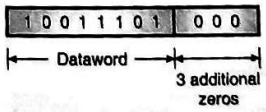
Generator polynomial :

$x^3 + 1 = x^3 + 0x^2 + 0x + 1 = 1001 = n$

Step 1 : Obtain the dividend :

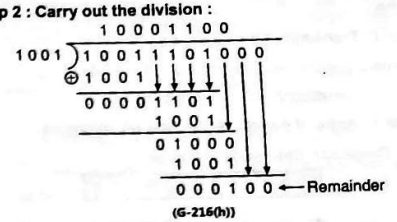
Dividend = Data word + 3 zeros.

- The dividend is as follows :



(G-216(g))

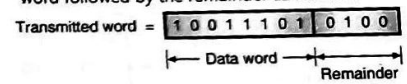
Step 2 : Carry out the division :



(G-216(h))

Step 3 : Obtain the actually transmitted bit stream :

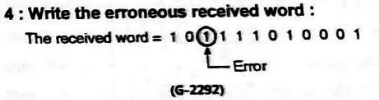
- The transmitted word is obtained by writing the data word followed by the remainder as follows :



(G-216(i))

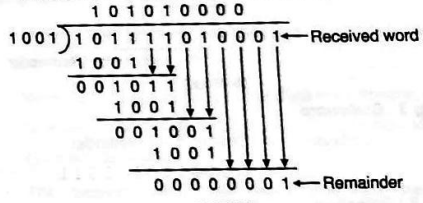
Error detection :

Step 4 : Write the erroneous received word :



(G-229Z)

- At the receiver, this word is divided by the same divider used at the transmitter i.e. 1001.



(G-216(j))

- A non zero remainder indicates that there is an error in the received codeword.

Ex. 3.8.5 : Consider an error detecting CRC with the generator 10101. Assume the CRC bits follows the data bits in any transmission :

1. Compute the transmitted bit sequence for the data bit sequence 01101101.
2. The string of bits 110011001100 is received. Is it acceptable, and if so what is the data bit sequence.

Dec. 03, 10 Marks

Soln. :

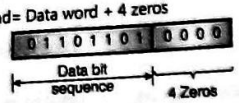
Part I : Transmitted bit sequence :

Given : Data bit sequence : 01101101

Generator : 10101

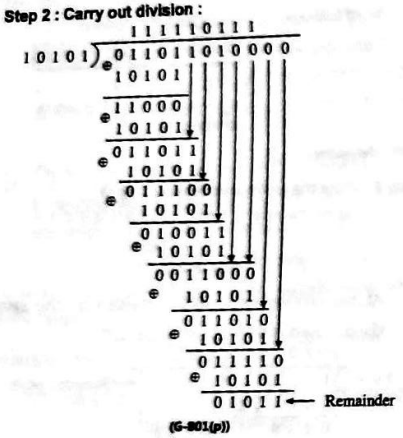
Step 1 : Append 4 zeros to the data bit sequence :

Dividend= Data word + 4 zeros



(G-801(o))

Step 2 : Carry out division :



(G-801(p))

Step 3 : Codeword :

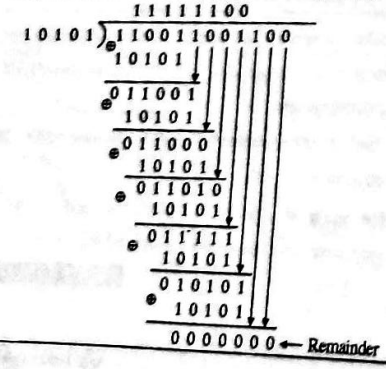
Codeword = Dividend + Remainder

= 01101101 1011

Part II : Detection

Received word : 110011001100

Carry out the division as follows : (G-801(q))



- Since the remainder is 0, the received codeword is acceptable and does not contain errors. So it is acceptable.

Received codeword = 11001100 1100

(G-801(r))

Ex. 3.8.6 : Consider an error detecting CRC with the generator 10101.

1. Compute the transmitted bit sequence for the data bit sequence 1101101.
2. The string of bits 110011001100 is received. Is acceptable, and if so what is the data bit sequence.

Dec. 05. 10 Marks

Soln. :

Part I : Transmitted bit sequence :

Given : Data bit sequence : 1101101

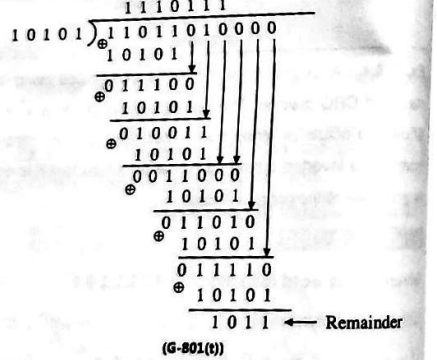
Generator : 10101

Step 1 : Obtain the dividend :

Dividend = Data word followed by 4 zeros.

= 1101101 0000

Step 2 : Carry out division :



(G-801(t))

Step 3 : Transmitted bit sequence :

- The transmitted bit sequence is obtained by adding remainder to the dividend.

∴ Transmitted code word = 1101101 1011

Data bits Remainder

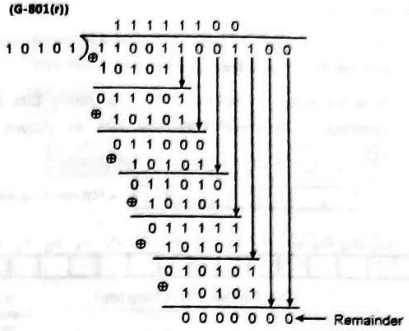
(G-801(u))

Part II : Decoding

- The received bit sequence is 1100 1100 1100. Divide it by the same generator used in part I.

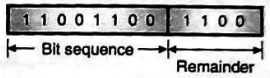
Received codeword = 11001100 1100

(G-801(r))



(G-801(v))

- Since the remainder is zero, there are no errors in the received bit sequence. Hence it is acceptable. The received sequence is,



(G-801(w))

∴ Received bit sequence = 11001100 ...Ans.

3.9 Error Correction : MU : Dec. 12

University Questions

Q. 1 Explain the error detection and error correction algorithms. **(Dec. 12, 10 Marks)**

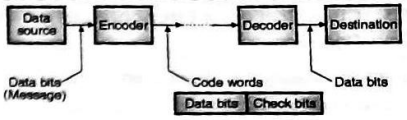
- We are going to discuss two completely different approaches for the error control.
- They are :
  1. Forward error correction (FEC).
  2. Automatic request for retransmission (ARQ).

The ARQ technique :

- In the ARQ system, the receiver can request for the retransmission of the complete or a part of message if it finds some error in the received message.
- This needs an additional channel called feedback channel to send the receiver's request for retransmission.

Error correction using redundant bits :

- In the error correction techniques, codes are generated at transmitter by adding a group of parity bits or check bits as shown in Fig. 3.9.1.

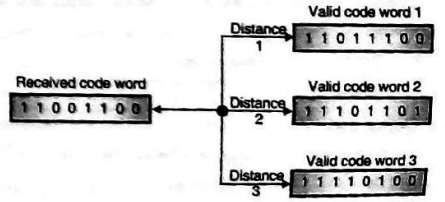


(L-306) **Fig. 3.9.1 : Error correction technique**

- The source generates the data (message) in the form of binary symbols.
- The encoder accepts these bits and adds the check (parity) bits to them to produce the code words.
- These code words are transmitted towards the receiver.
- The check bits are used by the decoder to detect and correct the errors.
- The encoder of Fig. 3.9.1, adds the check bits to the data bits, **according to a prescribed rule.** This rule will be dependent on the type of code being used.
- The decoder separates out the data and check bits. It uses the parity bits to detect and correct errors if they are present in the received code words.
- The data bits are then passed on to the destination.

FEC (Forward Error Correction) :

- In FEC the receiver searches for the most likely correct code word.
- When an error is detected, the distance between the received invalid code word and all the possible valid code words is obtained.
- The nearest valid code word (the one having minimum distance) is the most likely the correct version of the received code word as shown in Fig. 3.9.2.



(L-307) **Fig. 3.9.2 : Concept of FEC**

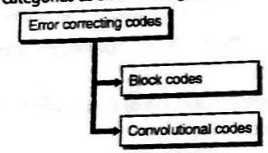
- In Fig. 3.9.2, the valid code word 1 has the minimum distance (1), hence it is the most likely correct code word.

3.9.1 Classification of Error-correcting Codes :

- We can classify the error correcting codes on the basis of different parameters.
- One way of classifying codes is :
  1. Linear codes 2. Non-linear codes.

Classification Based on the Functioning :

- The error correcting codes are broadly classified into two categories as shown in Fig. 3.9.3.



(E-173(a)) **Fig. 3.9.3**

- The major difference between the two is that the block codes don't need memory whereas convolutional codes do need memory.

3.9.2 Linear Block Codes :

- Block codes can be of two types : Linear block codes and nonlinear block codes.
- Almost all the block codes used today are **linear block codes.**
- The non linear block codes are not used widely as their analysis and implementation is difficult.

Definition :

- A linear block code is defined as that code in which the exclusive OR (Modulo – 2 addition) of two valid codewords would produce an another valid codeword.

3.9.3 Some Linear Block Codes :

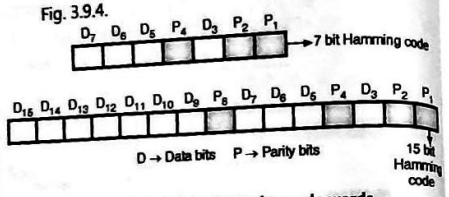
- Some of the important linear block codes are as follows :
  1. Simple parity check codes.
  2. Hamming codes.
- Out of these the parity checking codes are designed for error detection purpose whereas hamming codes are designed for error correction.

3.9.4 Hamming Codes :

- Now let us discuss the other category of codes i.e. the error correcting codes known as the Hamming codes.
- Hamming codes are linear block codes.

Hamming code structure :

- Hamming code is basically a linear block code named after its inventor.
- It is an error correcting code. The parity bits are inserted in between the data bits as shown in Fig. 3.9.4.

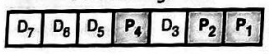


(G-1946) **Fig. 3.9.4 : Hamming code words**

- The 7-bit Hamming code is used commonly, but the concept can be extended to any number of bits.
- Note that the parity bits are inserted at each 2^n bit where n = 0, 1, 2, 3 ..... .
- Thus P_1 is at $2^0 = 1$, i.e. at first bit, P_2 is at $2^2 = 2$ P_4 is as 2^2 and P_8 is at 2^3 as shown in Fig. 3.9.4.

7-Bit Hamming Code :

1. *A scientist named R.W. Hamming developed a coding system which was easy to implement. Assuming that four data bits are to be transmitted, he suggested a code word pattern shown in Fig. 3.9.5.*



$D_7 D_6 D_5 D_3$ → Data bits
$P_4 P_2 P_1$ → Parity bits

(G-1947) **Fig. 3.9.5 : Code word pattern for Hamming code**

2. *The D bits in Fig. 3.9.5 are data bits, whereas P bits are parity bits. The parity bits P_1, P_2, P_4 are adjusted in a particular way as explained below.*

Minimum number of parity bits :

- *Table 3.9.1(a) gives a listing of minimum number of parity bits needed for various ranges of "m" information bits.*

Table 3.9.1(a) : Number of parity bits to be used

Number of information bits	Number of parity bits
2 to 4	3
5 to 11	4
12 to 26	5
27 to 57	6
58 to 120	7

Deciding the values of parity bits :

- Table 3.9.1(b) indicates which bit positions are associated with each parity bit in order to establish required parity (even or odd) over the selected bits positions.

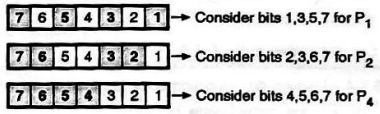
Table 3.9.1(b)

Parity Bit	Bits to be checked
P_1	1,3,5,7,9,11,13,15,.....
P_2	2,3,6,7,10,11,14,15, ...
P_4	4,5,6,7,12,13,14,15, ...
P_8	8,9,10,11,12,13,14,15, ...

Selection of parity bits :

Selection of P_1 :

- P_1 is adjusted to 0 or 1 so as to establish even parity over bits 1, 3, 5 and 7 i.e. P_1, D_3, D_5 and D_7.



(G-2291)

Selection of P_2 :

- P_2 is adjusted to 0 or 1 so as to set even parity over bits 2, 3, 6 and 7 (P_2, D_3, D_6 and D_7).

Selection of P_4 :

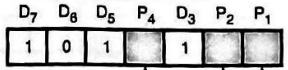
- P_4 is adjusted to 0 or 1 so as to set even parity over bits 4, 5, 6 and 7 (P_4, D_5, D_6 and D_7).
- The selection of parity bits will be clear after solving the following example.

Ex. 3.9.1 : A bit word 1 0 1 1 is to be transmitted. Construct the even parity seven-bit Hamming code for this data.

Soln. :

Step 1 : The codeword format :

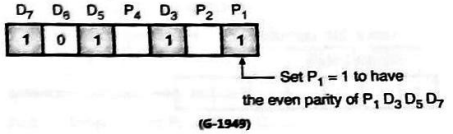
- The seven bit Hamming code format is shown in Fig. P. 3.9.1.
- Given bit word = 1 0 1 1



(G-1948) **Fig. P. 3.9.1**

Step 2 : Decide P_1 :

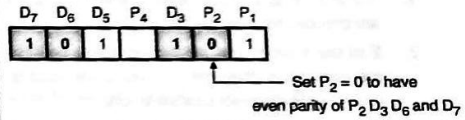
- P_1 sets the parity of bits P_1, D_3, D_5 and D_7. As D_7, D_5, D_3 = 1 1 1 we have to set P_1 = 1 in order to have the even parity.



(G-1949)

Step 3 : Decide P_2 :

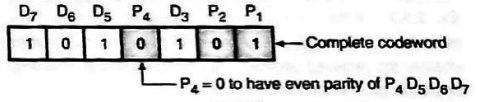
- P_2 is set to have the even parity of P_2 D_3 D_6 and D_7. But D_3 D_6 D_7 = 1 0 1 hence set P_2 = 0.



(G-1950)

Step 4 : Decide P_4 :

- P_4 is set to have the even parity of P_4 D_5 D_6 and D_7. But D_5 D_6 D_7 = 1 0 1, hence set P_4 = 0.

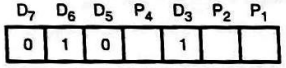


(G-1951)

Ex. 3.9.2 : Encode the data bits 0 1 0 1 into a seven bit even parity Hamming code.

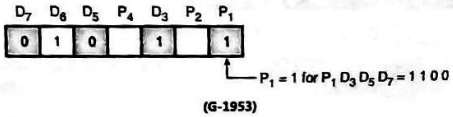
Soln. :

Step 1 :



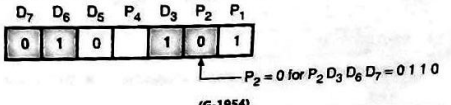
(G-1952)

Step 2 : Select P_1 for P_1 D_3 D_5 D_7 :



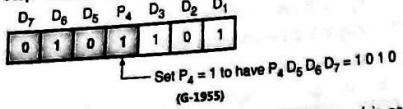
(G-1953)

Step 3 : Select P_2 for P_2 D_3 D_6 D_7 :



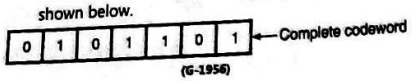
(G-1954)

Step 4 : Select P_4 :



(G-1955)

- Hence the complete 7-bit Hamming codeword is as shown below.



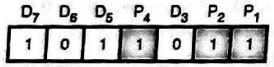
(G-1956)

Detection and correction of errors :

1. The Hamming coded data is now transmitted. At the receiver it is decoded to get the data back.
2. The bits (1, 3, 5, 7) , (2, 3, 6, 7) and (4, 5, 6, 7) are checked for even parity.
3. If all the 4-bit groups mentioned above possess the even parity then the received code word is correct i.e. it does not contain errors.
4. But if the parity is not even (i.e. it is odd) then error exists. Such an error can be located by forming a three bit number out of the three parity checks. This process becomes clear by solving the example given below.

Ex. 3.9.3 : If the 7-bit Hamming codeword received by a receiver is 1 0 1 1 0 1 1. Assuming the even parity state whether the received codeword is correct or wrong. If wrong, locate the bit in error.

Soln. :



(G-1957)

Step 1 : Analyze bits 4, 5, 6 and 7 :

$P_4\ D_5\ D_6\ D_7 = 1\ 1\ 0\ 1 \rightarrow$ Odd parity.

$\therefore$ error exists here.

$\therefore$ Put $P_4 = 1$ in the 4's position of the error word.

Step 2 : Analyze bits 2, 3, 6 and 7 :

$\therefore P_2\ D_3\ D_6\ D_7 = 1\ 0\ 0\ 1 \rightarrow$ Even parity so no error.

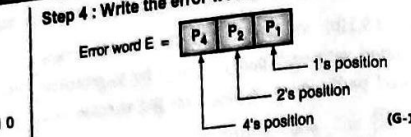
Hence put $P_2 = 0$ in the 2's position of the error word.

Step 3 : Check the bits 1, 3, 5, 7 :

$\therefore P_1\ D_3\ D_5\ D_7 = 1\ 0\ 1\ 1 \rightarrow$ Odd parity so error exists.

Hence put $P_1 = 1$ in the 1's position of the error word.

Step 4 : Write the error word :

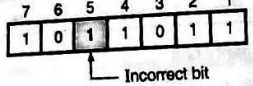


(G-1958)

- Substituting the values of P_4 , P_2 and P_1 obtained in steps 1, 2 and 3 we get,

$E = (5)_{10}$ (G-1959)

- Hence bit 5 of the transmitted codeword is in error.



(G-1960)

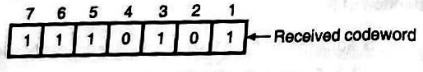
Step 5 : Correct the error :

- Invert the incorrect bit to obtain the correct codeword as follows :

Correct codeword = [1 0 0 1 0 1 1]

Ex. 3.9.4 : A seven bit Hamming code is received as 1 1 1 0 1 0 1. What is the correct code ?

Soln. :



(G-1961)

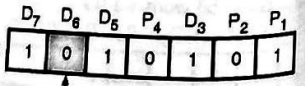
1. Check bits 4, 5, 6, 7 → Odd parity, hence error $\therefore P_4 = 1$
2. Check bits 2, 3, 6, 7 → Odd parity, hence error $\therefore P_2 = 1$
3. Check bits 1, 3, 5, 7 → Even parity, hence no error $\therefore P_1 = 0$

$\therefore$ Error word E = 1 1 0 ($P_4\ P_2\ P_1$)

(G-1968)

4. Decimal equivalent of $E = 1\ 1\ 0 = (6)_{10}$

$\therefore$ 6th bit in the received codeword is incorrect. So invert it



(G-1962)

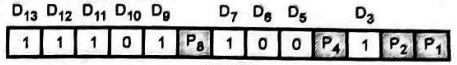
Ex. 3.9.5 : Generate the Hamming code for the data 111011001 with even parity.

Soln. :

Data number : 1 1 1 0 1 1 0 0 1

Step 1 :

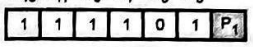
- Number of message bits is 9. So we need to add 4 parity bits in the codeword.
- The parity bits will be at the positions 1, 2, 4, and 8 as shown below :



(G-218(a))

Step 2 : Select P_1 for $P_1\ D_3\ D_5\ D_7\ D_9\ D_{11}\ D_{13}$:

- Parity needs to be even parity



(G-218(b))

- For even parity P_1 should be 1. $\therefore P_1 = 1$

Step 3 : Select P_2 :

- To select P_2 we have to consider the bits in positions 2, 3, 6, 7, 10 and 11

$\therefore\ 1\ 0\ 1\ 0\ 1\ P_2 \rightarrow P_2 = 1$ $\therefore P_2 = 1$

Step 4 : Select P_4 :

- For P_4, we have to consider the bits in the following positions 4, 5, 6, 7, 12, 13 and select the value of P_4 for even parity.

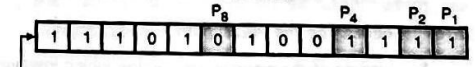
$\therefore\ 1\ 1\ 1\ 0\ 0\ P_4$ $\therefore P_4 = 1$

Step 5 : Select P_8 :

- To select P_8, consider the bit in following positions 8, 9, 10, 11, 12, 13 and select P_8 for even parity.

$\therefore\ 1\ 1\ 1\ 0\ 1\ P_8$ $\therefore P_8 = 0$

- So the codeword is as follows :



(G-218(c))

Ex. 3.9.6 : An 8-bit byte with binary value 10101111 is to be encoded using an even-parity Hamming code. What is the binary value after encoding ?

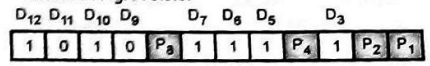
Dec. 10, 5 Marks, Dec. 11, May 12, 10 Marks

Soln. :

Given : Data number : 10101111

Step 1 :

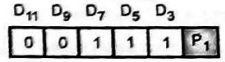
- Number of message bits is 8. So we need to add 4 parity bits in the codeword.
- The parity bits will be at the positions 1, 2, 4 and 8 as shown in Fig. P. 3.9.6.



(G-2840)**Fig. P. 3.9.6**

Step 2 : Select P_1 for $P_1\ D_3\ D_5\ D_7\ D_9\ D_{11}$:

- Parity needs to be even parity.

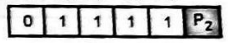


(G-2841)**Fig. P. 3.9.6(a)**

- For even parity P_1 should be 1.

$\therefore\ P_1 = 1$

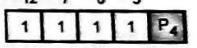
Step 3 : Select P_2 for $P_2\ D_3\ D_6\ D_7\ D_{10}\ D_{11}$:



(G-2842)**Fig. P. 3.9.6(b)**

$\therefore\ P_2 = 0$

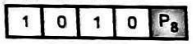
Step 4 : Select P_4 for $P_4\ D_5\ D_6\ D_7\ D_{12}$:



(G-2843)**Fig. P. 3.9.6(c)**

$\therefore\ P_4 = 0$

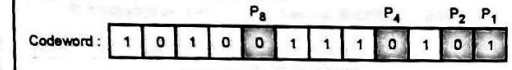
Step 5 : Select P_8 for $P_8\ D_9\ D_{10}\ D_{11}\ D_{12}$:



(G-2844)**Fig. P. 3.9.6(d)**

$\therefore\ P_8 = 0$

- So codeword is as follows,

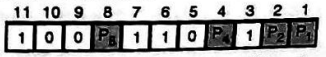


(G-2845) **Fig. P. 3.9.6(e)**

Ex. 3.9.7 : Compute the Hamming code for the data - 1001101. Dec. 13, 5 Marks

Soln. :

Step 1 : Codeword format :



(G-2278) **Fig. P. 3.9.7 : Codeword format**

Step 2 : Find : $P_1P_2P_4P_8$:

- Assume even parity.

1. P_1 :

- Consider bits 1,3,5,7,9,11 They are,
  10101 P_1 ∴ For even parity $P_1 = 1$

2. P_2 :

- Consider bits 2,3,6,7,10,11 They are,
  10111 P_2 ∴ For even parity $P_2 = 0$

3. P_4 :

- Consider bits 4,5,6,7 They are,
  110 P_4 ∴ For even parity $P_4 = 0$

4. P_8 :

- Consider bits 8,9,10,11 They are,
  100 P_8 ∴ For even parity $P_8 = 1$

Step 3 : Write the codeword :

Code word = 1 0 0 1 1 1 0 0 1 0 1

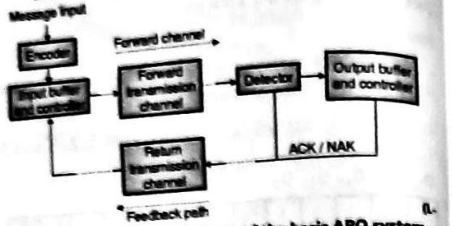
(6-3279)

3.10 ARQ Technique :

- There are two basic systems of error detection and correction.
- The first on being the forward error correction (FEC) system and the second one is the automatic repeat request (ARQ) system.
- In the ARQ system of error control, when an error is detected, a request is made for the retransmission of that signal.
- Therefore a feedback channel is required for sending the request for retransmission.
- The ARQ systems differ from the FEC systems in three important respects. They are as follows :
  1. In ARQ system less number of check bits (parity bits) are required to be sent. This will increase the (k/n) ratio for an (n,k) block code if transmitted using the ARQ system.
  2. A return transmission path and additional hardware in order to implement repeat transmission of codewords will be needed.
  3. The bit rate of forward transmission must make allowance for the backward repeat transmission.

Basic ARQ system :

- The block diagram of the basic ARQ system is as shown in Fig. 3.10.1.



372) **Fig. 3.10.1: Block diagram of the basic ARQ system**

Operation of ARQ system :

- The encoder produces codewords for each message signal at its input.
- Each codeword at the encoder output is stored temporarily and transmitted over the forward transmission channel.
- At the destination a decoder will decode the code words and look for errors.
- The decoder will output a "positive acknowledgment" (ACK) if no errors are detected and it will output a negative acknowledgment (NAK) if errors are detected.
- On receiving a negative acknowledgment (NAK) signal via the return transmission path the "controller" will retransmit the appropriate word from the words stored by the input buffer.
- A particular word may be retransmitted only once or it may be retransmitted twice or more number of times.
- The output controller and buffer on the receiver side assemble the output bit stream from the code words accepted by the decoder.

Error probability on the return path :

- The bit rate of the return transmission which involves the return transmission of ACK/NAK signal is low as compared to the bit rate of the forward transmission.
- Therefore the error probability of the return transmission is negligibly small.

Types of ARQ system :

- The three types of ARQ systems are :
  1. Stop-and-wait ARQ system

2. Go back n ARQ and 3. Selective repeat ARQ.

Note : Error control in the data link layer is based on the principle of request for automatic retransmission (ARQ) of the missing, lost or damaged frames.

3.11 Flow Control :

MU Dec 09, Dec 11, Dec 14

University Questions

Q. 1 Explain framing, flow and error control in data link layer. **(Dec. 09, Dec. 11, 10 Marks)**

Q. 2 Why is flow control needed ? What are the mechanisms ? Explain how the Go-Back-N and Selective Repeat ARQ differ from each other. **(Dec. 14, 10 Marks)**

- This is another important design issue related to the data link layer.
- In flow control the problem to be handled is what to do with the sender computer that wants to send data at a faster rate than the capacity of the receiver to receive them.
- This happens when the sender is using a faster computer than the receiver.
- The data sent at a very fast rate will completely overwhelm the receiver.
- The receiver will keep losing some of the frames simply because they are arriving too quickly.
- The solution to this problem is to introduce the **flow control.**
- The flow control will control the rate of frame transmission to a value which can be handled by the receiver.
- It requires some kind of a feedback mechanism from the receiver to the sender, so as to adjust the sending rate automatically.
- We are going to discuss some flow control techniques based on this principle.
- It is a set of procedures that tells the sender how much data it can transmit before it must wait for an acknowledgement from the receiver, otherwise there will be overflow of data.
- The data flow should not be so fast that the receiver is over-whelmed.
- The speed of processing of any receiving device is a limited and it also has a limited amount of memory storage space, for storing the incoming data.
- There has to be some system, for reverse communication from the receiver to transmitter.
- The receiver can tell the transmitter about adjusting the data flow rate to suit its speed or even stop temporarily.
- As the rate of processing at the receiver is generally slower than the rate of transmission.
- Each receiver has a finite memory called **buffer.**
- The incoming data is first stored in the buffer and then sequentially processed.
- If the buffer begins to fill up, the receiver must be able to tell the sender to stop transmission until the buffer gets empty.
- Similarly the transmitter also has a buffer for storing the bits if the transmission is stopped.

Note : Flow control can be defined as a set of procedures which are used for limiting the amount of data a transmitter can send before waiting for acknowledgement.

3.12 Elementary Data Link Protocols :

- In this section we are going to discuss some elementary data link layer protocols.

3.12.1 An Unrestricted Simplex Protocol :

- This protocol is the simplest possible protocol.
- The transmission of data takes place in only one direction. So it is a simplex (unidirectional) protocol.
- It is assumed that the network layers of sender and receiver are always ready.
- It is also assumed that we can ignore the processing time and the buffer space available infinite.
- The communication channel is imagined to be noise free so it does not damage or lose any frames.
- All this is highly unrealistic. This protocol is also called as "utopia".
- This protocol consists of two distinct procedures, namely a sender and a receiver.

- They run in the data link layers of their respective machines.
- No sequence numbers or acknowledgements are used.

3.12.2 A Simplex Stop and Wait Protocol :

MU : Dec. 03, May 07, Dec. 08, Dec. 09

University Questions

Q. 1 Explain stop and wait sliding window protocol. **(Dec. 03, 10 Marks)**

Q. 2 Explain stop and wait sliding window protocols with suitable examples. **(May 07, Dec. 08, Dec. 09, 10 Marks)**

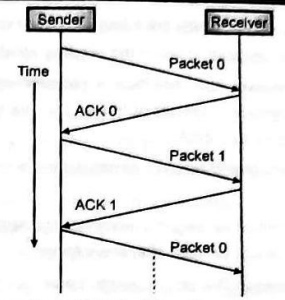
- The most unrealistic restriction in the previous protocol is the assumption that the receiving network layer can process the data with zero processing time.
- In the simplex stop and wait protocol it is assumed that a finite processing time is essential.
- However like the first protocol, the communication channel is assumed to be noise free and the communication is simplex i.e. only in one direction at a given time.
- This protocol deals with an important problem i.e. how to prevent the sender from flooding the receiver due to the data rates faster than processing speed of the receiver.
- In this protocol, a small dummy frame is sent back from the receiver to the transmitter to indicate that it can send the next frame. The small dummy frame is called as acknowledgement.
- The transmitter sends one frame and then waits for the dummy frame called acknowledgement.
- Once the acknowledgement is received, it sends the next frame and waits for the acknowledgement. Hence this protocol is known as **stop and wait** protocol.
- The best thing about this protocol is that the incoming frame is always an acknowledgement. It need not be even checked.

3.12.3 A Simplex Protocol for Noisy Channel :

- This is the third protocol in which we go one step ahead and assume that the communication channel is noisy and can introduce errors in the data travelling over it.
- The channel noise can either damage the frames or they may get lost completely.
- In this protocol, the sender waits for a positive acknowledgement before advancing to the next data item.
- There is a timer set at the sender when a frame is sent. If the sender times out it will resend the same frame again.
- So it is called as PAR (Positive acknowledgement with retransmission) or Automatic Repeat Request (ARQ) type protocol.
- If a frame is badly damaged or lost then the sender would retransmit it.
- Note that due to retransmission (time out or any other reason), there is always a possibility of duplication of frames at the receiver.
- To avoid this, the sender puts a sequence number in the header of each frame it sends.
- The receiver can check the sequence number of each arriving frame to check for the possible duplicate frame.
- If a frame is duplicated then receiver will discard it.
- The operation can be divided into two modes :
  1. Normal operation and
  2. Time out.

1. Normal operation :

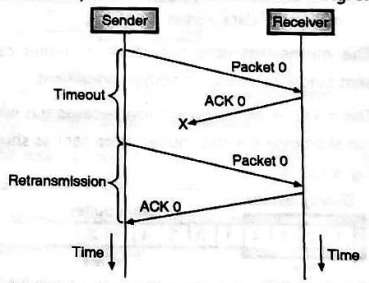
- After transmitting one frame, the sender waits for an acknowledgement (ACK) from the receiver before transmitting the next one.
- In this way, the sender can recognize that the previous packet is transmitted successfully and we could say "stop and wait" guarantees reliable transfer between nodes.
- To support this feature, the sender keeps a record of each frame it sends.
- Also, to avoid confusion caused by delayed or duplicated ACKs, "stop-and-wait" sends each packet with unique sequence numbers and receives that numbers in each ACKs.



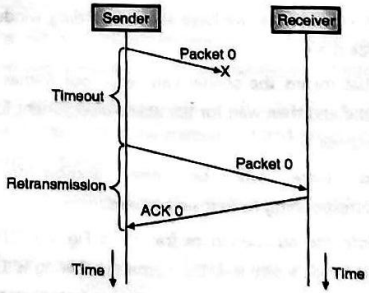
(G-220)**Fig. 3.12.1 : Positive acknowledgement with retransmission**

2. Time out :

- If the sender does not receive **ACK** for previous sent frame after a certain period of time, the sender times out and retransmit that frame again.
- There are two cases when the sender does not receive ACK; one is when the ACK is lost and the other is when the frame itself is not received i.e. it got lost.
- These two possible cases are illustrated in Fig. 3.12.2.



(a) Sender does not receive acknowledgement



(b) Frame is lost

(G-221) **Fig. 3.12.2 : Timeout and retransmission**

- To support this feature, the sender keeps timer for each frame.
- We have already discussed that a timer is introduced in the data link layer.

3.12.4 Piggybacking :

MU : Dec. 07, Dec. 19

University Questions

Q. 1 Describe in brief piggybacking. **(Dec. 07, 4 Marks)**

Q. 2 Describe in brief the concept of piggybacking. **(Dec. 19, 5 Marks)**

- In all the practical situations, the transmission of data needs to be bi-directional. This is called as full-duplex transmission.
- One way of achieving full duplex transmission is to have two separate channels one for forward data transmission and the other for reverse data transfer (for acknowledgements).
- But this will waste the bandwidth of the reverse channel almost entirely.
- A better solution would be to use each channel (forward and reverse) to transmit frames bothways, with both channels having the same capacity.
- Let A and B be the users. Then the data frames from A to B are intermixed with the acknowledgements from A to B.
- By checking the kind field in the header of the received frame the received frame can be identified as either data frame or acknowledgement.
- One more improvement can be made. When a data frame arrives, the receiver waits, does not send the control frame (acknowledgement) back immediately.
- The receiver waits until its network layer passes in the next data packet.
- The acknowledgement is then attached to this outgoing data frame.
- Thus the acknowledgement travels alongwith next data frame.
- This technique in which the outgoing acknowledgement is delayed temporarily is called as **piggybacking**.

Advantage of piggybacking :

- The major advantage of piggybacking is better use of available channel bandwidth.
- This happens because an acknowledgement frame need not be sent separately.

Disadvantages :

1. The disadvantage of piggybacking is the additional complexity.
2. If the data link layer waits too long before transmitting acknowledgement, then retransmission of frame would take place.

3.13 Sliding Window Protocols :

MU - Dec. 03, May 05, May 07, Dec. 08, Dec. 09

University Questions

Q. 1 Explain stop and wait and sliding window protocol. **(Dec. 03, 10 Marks)**

Q. 2 Explain stop and wait and sliding window protocols with suitable examples. **(May 05, May 07, Dec. 08, Dec. 09, 10 Marks)**

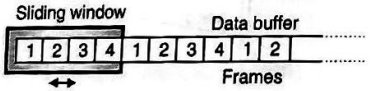
- The next three protocols are more robust and bi-directional protocols.
- All these protocols are special type of protocol called **Sliding Window Protocols.**
- They show a different performance in terms of their efficiency, complexity and buffer requirements.

Sequence number :

- One of the important features of all the sliding window protocols is that each outbound frame contains a sequence number, ranging from 0 to some maximum value.
- The maximum value is generally equal to $(2^n - 1)$. The value of n can be arbitrary.

Sliding windows :

- Sliding windows are basically the imaginary boxes at the transmitter and receiver.
- This window holds the frames at the transmitting as well as receiving ends and provides the upper limit on the number of frames that can be transmitted before acknowledgement is obtained.
- So in short we can say that, at any instant of time, the sender maintains a set of sequence numbers corresponding to the frames it is permitted to send.
- These frames which are being permitted to sent are said to be residing inside the **sending window.**
- The receiver also maintains a **receiver window.** It corresponds to the set of frames that the receiver is permitted to accept.
- The sender and receiver windows can be of different sizes.
- The positive or negative acknowledgement (ACK or NAK) should be used after every frame.
- That means the sender sends frame, waits for the acknowledgement and sends the next frame or retransmits the original one, only after receiving either positive or negative acknowledgement from the receiver.
- In order to improve the efficiency, the sender sends multiple frames at time, the receiver checks the CRC of all the frames one by one and sends one acknowledgement for all the frames.
- This is the principle of operation of sliding window technique.
- In this technique, an imaginary window consisting of "n" number of data frames is defined.
- This means that upto n number of frames can be sent before receiving an acknowledgement.
- This is known as sliding window because this window can slide over the data buffer to be sent as shown in Fig. 3.13.1(a).

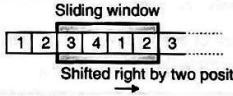


(G-222)Fig. 3.13.1(a) : Sliding window

- In Fig. 3.13.1(a) we have shown a sliding window of size n = 4.
- That means the sender can send four frames, at a time and then wait for the acknowledgement for the receiver.
- So there will be one acknowledgement corresponding to four sent frames.
- Note the numbering of frames in Fig. 3.13.1(a). As the window size is 4, the frame numbering is 1, 2, 3, 4 then again 1, 2, 3, ... the maximum frame number is restricted to n.

Sender and receiver sliding windows :

- The sender as well as the receiver maintain their own sliding windows.
- The sender sends the number of frames allowed by the size of its own sliding window and then waits for an acknowledgement from the receiver.
- The receiver sends an acknowledgement which includes the number of the next frame that the sender should send.
- For example if the sender has sent frames 1 and 2 to the receiver and if receiver receives them correctly, then the acknowledgement sent by the receiver will include number-3 indicating the sender to send frame number-3.
- Now if the sender transmits the first 4 frames as per the size of its window and receives an acknowledgement for the first two frames, then the sender will slide its window two frames to the right as shown in Fig. 3.13.1(b) and sends 5th and 6th frames (i.e. frames 1 and 2 of the next lot).

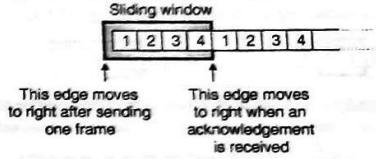


(G-223)Fig. 3.13.1(b) : Illustration of sliding window mechanism

- The receiver now has four frames again, so it checks frames 3, 4, 1, 2 by checking their CRC.
- If it finds frame 3 faulty then it will send an acknowledgement which includes number 3.
- The sender will send 4-frames staring from frame-3 onwards.
- The sliding window mechanism thus uses two buffers and one window so as to exercise the flow control.
- The application program on the sender side will create the data to be transmitted and loads into the sender's buffer.
- Then the sender's sliding window is imposed on this buffer.
- These frames are then sent till all the frames have been sent.
- The receiver receives these data frames and carriers out checks such as CRC, missing or duplicate frames etc. and stores the correct frames in the receiver buffer.
- The application program at the receiver then the takes this data.

Movement of sender's window :

- Fig. 3.13.1(c) shows the sender's window.



(G-224)Fig. 3.13.1(c) : Sender sliding window

- If the senders window size is 4 and frames 1 and 2 are sent but acknowledgement has not been received so far, then as shown in Fig. 3.13.1(d), the sender's windows will only contain two frames i.e. 3 and 4.

(G-225)Fig. 3.13.1(d) : Sender's window after sending first two frames but no acknowledgement

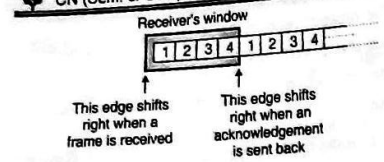
- Now if the sender receives acknowledgement bearing number 3 then it understands that the receiver has correctly received frames 1 and 2.
- The senders window now expands and includes the next two frames as shown in Fig. 3.13.1(e).

(G-226)Fig. 3.13.1(e) : Sender's window after receiving acknowledgement bearing number-3

- In this way the left edge of senders window will shift right when the data frames are sent and the right edge of the senders window will shift right when the acknowledgement is received.

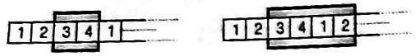
Movement of receiver's windows :

- Fig. 3.13.1(f) shows the receiver's window. Its left edge shifts right on receiving each data frame, where as its right edge shifts right when an acknowledgement is sent.



(G-227)**Fig. 3.13.1(f) : Receiver's sliding window**

- If we take the same example that we discussed for the sender's window then the position of receivers windows are as shown in Fig. 3.13.1(g) and (h).



(g) Two frames (1 and 2) received but no acknowledgement sent **(h) After sending the acknowledgement**

(G-228)**Fig. 3.13.1 : Movement of receiver window**

Ex. 3.13.1 : Two neighbouring nodes A and B uses sliding window protocol with 3 bit sequence number. As the ARQ mechanism Go back N is used with window size of 4. Assume A is transmitting and B is receiving show window position for the following events :

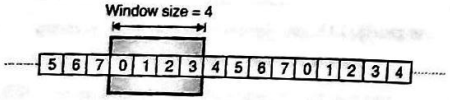
1. Before A send any frame.
2. After A send frame 0, 1, 2 and receive ACK (acknowledgement) from B for 0 and 1.

Soln. :

- The number of sequence number bits = m = 3.

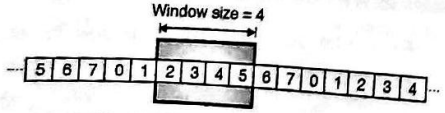
  ∴ The sequence numbers will be 0, 1, 2, 3 ..., 6, 7. We can repeat these numbers. So the sequence will be,

  0, 1, 2, 3, 4, 5, 6, 7, 0, 1, 2, 3, 4, ...
- The size of the window is 4.
- Fig. P. 3.13.1(a) shows the sender window (at A) before sending any frame.



(G-229) **Fig. P. 3.13.1(a) : Before A sends any frame**

- Fig. P. 3.13.1(b) shows that the window slides 2 positions because acknowledgement for frames 0 and 1 have been received.



(G-230) **Fig. P. 3.13.1(b) : After sliding two frames**

3.13.1 A One Bit Sliding Window Protocol (Stop and Wait ARQ) :

MU : May 04, May 05, Dec. 05, May 10

University Questions

Q. 1 Explain a one - bit sliding window protocol in detail. **(May 04, May 10, 10 Marks)**

Q. 2 Explain stop and wait and sliding window protocol with example and suitable diagrams. **(May 05, 10 Marks)**

Q. 3 Explain n-bit sliding window protocol. **(Dec. 05, 6 Marks)**

- This protocol is called one bit protocol because the maximum window size here i.e. n is equal to 1.
- It uses the stop-and-wait technique which we have discussed earlier. The sender sends one frame and waits to get its acknowledgement.
- The sender transmits its next frame only after receiving the acknowledgement for the earlier frame.
- So one bit sliding window protocol is also called as **stop and wait protocol.**
- The sequence of events taking place when a frame is transmitted and received is as follows :

1. The data link layer of the sending machine fetches the first packet from its network layer.
2. It builds the frame for it and sends it to receiver.
3. The receiver data link layer checks the received frame for duplication.
4. If ok, it passes the frame to its network layer.

(G-234)

The operation of protocol :

- The operation of this protocol is based on the ARQ (automatic repeat request) principle.
- So the sliding window protocols are also called as ARQ protocols.
- In this method the transmitter transmits one frame of data and waits for an acknowledgement from the receiver.

- If it receives a positive acknowledgement (ACK) it transmits the next frame.
- If it receives a negative acknowledgement (NAK) it retransmits the same frame.

Features added for retransmission :

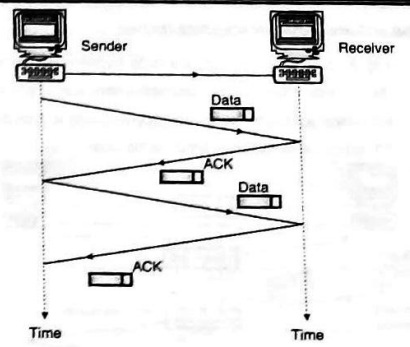
- For retransmission, four features are added to the basic flow control mechanism :
  1. The transmitter stores the copy of last frame transmitted until an acknowledgement for that frame is received from the destination.
  2. For distinctly identifying different types of frames both data and ACK frames are numbered alternately 0 and 1. The first data frame sent is numbered as 0. This frame is acknowledged by an ACK 1 frame. After receiving ACK1 the sender sends next data frame having a number 1.
  3. If an error occurs while transmission, the receiver sends a NAK frame back to the transmitter for retransmission of the corrupted frame. NAK frames which are not numbered tell the transmitter to retransmit the last frame transmitted.
  4. The transmitter has a timer to take care of the frame ACK which are lost. After a specified time if the transmitter does not receive a ACK or NAK frame it retransmits the last frame.

When is the retransmission necessary ?

- The retransmission of frame is essential under the following events :
  1. If the received frame is damaged.
  2. If the transmitted frame is lost.
  3. If the acknowledgement from the receiver is lost.
- Let us see the operation of the protocol under these circumstances one by one.

Operation under normal condition :

- Fig. 3.13.2 illustrates the protocol operation when everything is normal.

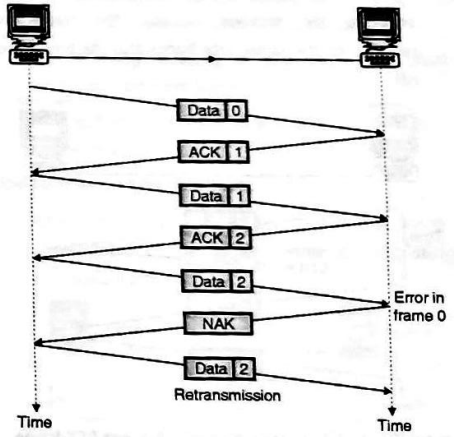


(G-235) **Fig. 3.13.2 : Stop and wait under normal condition**

- No frame is lost so retransmission is not necessary.

Stop and wait ARQ for damaged frame :

- As seen in Fig. 3.13.3(a) the transmitter transmits data frame numbered 0.

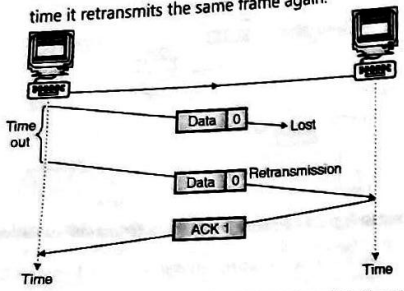


(G-236) **Fig. 3.13.3(a) : Stop and wait ARQ damaged frame**

- The receiver returns an ACK 1 indicating that the data frame numbered 0 is received without any error. The next data frame i.e. data 1 is sent.
- The corresponding acknowledgement ACK2 is received.
- The process goes on in this way, but if an error occurs the receiver sends a NAK requesting retransmission of the corrupted data frame (data 2).
- So the transmitter retransmits the data frame 2.

Stop and wait ARQ for lost data frame :

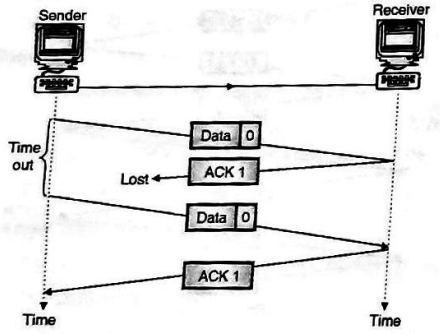
- Fig. 3.13.3(b) shows that if a data frame is lost and if the transmitter does not receive any type of acknowledgement from the receiver with a specified time it retransmits the same frame again.



(G-237) **Fig. 3.13.3(b) : Stop and wait ARQ, lost data frame**

Stop and wait ARQ for lost acknowledgement :

- Fig. 3.13.3(c) shows that if the acknowledgement sent by the receiver is lost, the transmitter retransmits the same data frame after its timer goes off.



(G-238) **Fig. 3.13.3(c) : Stop and wait ARQ, lost ACK frame**

- Stop and wait ARQ protocol becomes inefficient when the propagation delay is much greater than the time to transmit a frame. e.g. let us assume that we are transmitting frames that are 800 bits long over a channel that has a speed of 1 Mbps and let us also assume that the time taken for transmission of the frame and its acknowledgement is 30 mS.
- The number of bits that can be transmitted over this channel in 30 mS is equal to $30 \times 10^{-3} \times 1 \times 10^{6}$ = 30,000 bits.
- But in the stop-and-wait ARQ only 800 bits can be transmitted in this time period.
- This inefficiency is due to the fact that in stop and wait ARQ the transmitter waits, for an acknowledgement from the receiver before sending the next frame.
- The product of the bit rate and the delay that elapses before an action can take place is called the Delay-bandwidth product.
- The Delay-bandwidth product helps in measuring the lost opportunity in terms of transmitted bits.

Note : Stop-and Wait ARQ was used in IBM's Binary Synchronous Communications (Bisync) Protocol. It is also used in Xmodem, a popular file transfer protocol for modem.

Disadvantages of stop and wait protocol :

1. Problem with Stop-and-Wait protocol is that it is very inefficient. At any one moment, only one frame is in transition.
2. The sender will have to wait at least one round trip time before sending next. The waiting can be long for a slow network such as satellite link.

3.13.2 A Protocol using GO Back n :

MU : May 11, Dec. 11, Dec. 14, May 15, Dec. 17, Dec. 19

University Questions

Q. 1 Explain sliding window protocol using go-back N technique. **(May 11, Dec. 11, May 15, Dec. 17, 10 Marks)**

Q. 2 Why is flow control needed ? What are the mechanisms ? Explain how the Go-Back-N and Selective Repeat ARQ differ from each other. **(Dec. 14, 10 Marks)**

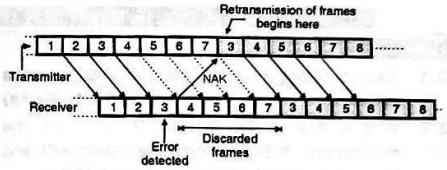
Q. 3 Explain the Go-back-N protocol. **(Dec. 19, 10 Marks)**

- In this stop and wait protocol it was assumed that the transmission time required for a frame to arrive at the receiver plus the transmission time for the acknowledgement to come back is negligible.
- But in some practical situations, this assumption is not correct.
- In the systems like satellite system the round trip time can be as long as 500 mS (propagation delay). This will reduce the efficiency of the protocol.

- Therefore an improved protocol known as **GO-Back-n ARQ** has been developed.
- It is a method used to overcome the inefficiency of the stop and wait ARQ by allowing the transmitter to continue sending enough frames so that the channel is kept busy while the transmitter waits for acknowledgements.
- In this method if one frame is damaged or lost, all frames are sent since the last frame acknowledged are retransmitted.

Principle of GO-back-n ARQ :

- Refer Fig. 3.13.4 to understand the principle of GO-Back-n ARQ.



(G-239) **Fig. 3.13.4 : Go back n ARQ system**

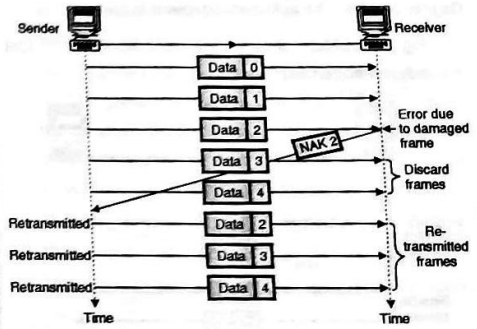
- The major difference between this and the previous system is that the sender does not wait for ACK signal for the transmission of next frame.
- It transmits the frames continuously as long as it does not receive the "NAK" signal. NAK is the negative acknowledgement signal sent by the receiver to the transmitter.
- When the receiver detects an error in the third frame as shown in Fig. 3.13.4, the receiver sends a NAK signal back to sender.
- But this signal takes some time to reach the transmitter. By that time the transmitter has transmitted frames upto frame 7.
- On reception of the NAK signal, the transmitter will retransmit all the frames from 3 onwards.
- The receiver discards all the frames it has received after 3 i.e. 3 to 7.
- It will then receive all the frames that are retransmitted by the transmitter.

Sources of error :

- The errors can get introduced, if the transmitted frames are damaged or lost or if the acknowledgement is lost.
- Let us consider the operation of this protocol under these conditions.

Operation when the frame is lost :

- This condition is illustrated in Fig. 3.13.5(a).

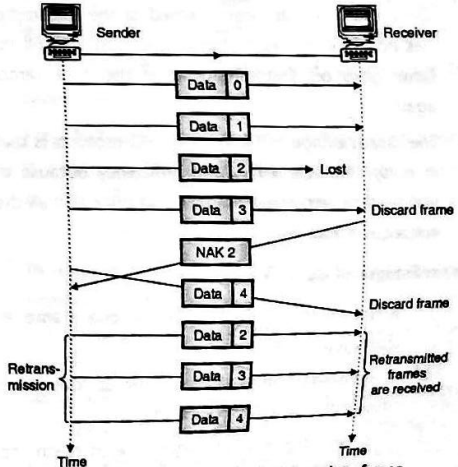


(G-240) **Fig. 3.13.5(a) : Go-back-n, damaged data frame**

- The second data frame is damaged, so the error is detected and receiver send NAK-2 signal back.
- On receiving this signal, the transmitter starts retransmission from frame 2.
- All the frames received after frame 2 are discarded by the receiver.

Operation when a frame is lost :

- As shown in Fig. 3.13.5(b) the case of lost frame is also treated in the same manner as that of the damaged frame.

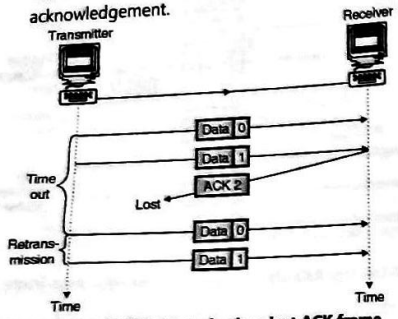


(G-241) **Fig. 3.13.5(b) : Go-back-n, lost data frame**

- The receiver, if it does not receive a particular data frame it sends a NAK to the transmitter and the transmitter retransmits all the frames sent since the last frame acknowledged.

Operation when the acknowledgement is lost :

- Fig. 3.13.5(c) shows the condition for lost acknowledgement.



(G-242)Fig. 3.13.5(c) : Go-back-n, lost ACK frame

- In case of go-back-n method the transmitter does not expect an acknowledgement after every data frame.
- It cannot use the absence of sequential ACK numbers to identify lost ACK or NAK frames, instead it uses a timer.
- The transmitter can send as many frames as the window allows before waiting for an acknowledgement.
- Once the limit has been reached or the transmitter has no more frames to transmit it must wait till the timer goes off and retransmit all the data frames again.
- The disadvantage of Go-back-n ARQ protocol is that in noisy channels it has poor efficiency because of the need to retransmit the frame in error and all the subsequent frames.

Disadvantages of Go back n :

1. It transmits all the frames if one frame is damaged or lost.
2. It transmits frames continuously as long as it does not receive the NAK signal.
3. The NAK signal takes some time to reach the sender. Till that time the sender has already sent some frames. All those will be retransmitted after receiving the NAK.
4. The error can get introduced if the NAK is lost.

3.13.3 Pipelining :

- In networking a new task is often started before the previous task has been completed. This is called pipelining.
- The principle of pipelining is not used in stop-and-wait ARQ but it is used in GO-Back-n ARQ and the selective repeat ARQ.
- Pipelining improves the efficiency of transmission.

3.13.4 Selective Repeat ARQ :

MU : May 13, Dec. 14, Dec. 15, May 17, May 19

University Questions

Q. 1 Explain sliding windows protocol with selective repeat. **(May 13, 10 Marks)**

Q. 2 Why is flow control needed ? What are the mechanisms ? Explain how the Go-Back-N and Selective Repeat ARQ differ from each other. **(Dec. 14, 10 Marks)**

Q. 3 What is the maximum window size allowed for selective repeat ARQ ? Explain why with appropriate scenario. **(Dec. 15, 10 Marks)**

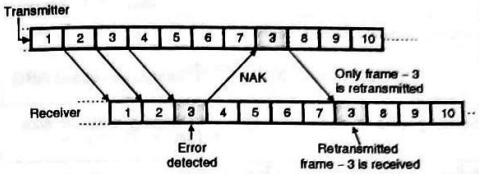
Q. 4 With the help of suitable example explain sliding window protocol with selective repeat. Compare its performance to sliding window with Go-back-n technique. **(May 17, 10 Marks)**

Q. 5 Explain design issues of Data Link layer. Explain Sliding Window protocol Selective Repeat. **(May 19, 10 Marks)**

- In this method only the specified damaged or lost frame is retransmitted.
- A selective repeat systems differs from the go-back-n method in the following ways :
  1. The receiver can do sorting of data frames and is also able to store frames received after it has sent the NAK until the damaged frame has been replaced.
  2. The transmitter has a searching mechanism that allows it to choose only those frames which are requested for retransmission.
  3. The window size in this method is less than or equal to $(n + 1)/2$, whereas in case of go-back-n it is $n - 1$.

Principle of Selective Repeat ARQ :

- The principle of operation of this protocol is illustrated in Fig. 3.13.6.



(G-243) Fig. 3.13.6 : Selective repeat ARQ system

- In this system as well, the transmitter does not wait for the ACK signal for the transmission of the next frame.
- It transmits the frames continuously till it receives the "NAK" signal from the receiver.
- The receiver sends the "NAK" signal back to the transmitter as soon as it detects an error in the received frame.
- For example the receiver detects an error in the third frame, as shown in Fig. 3.13.6.
- By the time this "NAK" signal reaches the transmitter, it had transmitted the frames upto 7 as shown in Fig. 3.13.6.
- On reception of "NAK" signal, the transmitter will retransmit only the frame-3 and then continues with the sequence 8, 9... as shown in Fig. 3.13.6.
- The frames 4, 5, 6 and 7 received by the receiver which do not contain any error are not discarded by the receiver.
- The receiver receives the retransmitted frames in between the regular frames.
- Therefore the receiver will have to maintain the frames sequentially.

Hence the selective repeat ARQ is the most efficient but the most complex protocol, of all the ARQ protocols.

- Thus in selective repeat ARQ only the frame which is damaged or lost is retransmitted by the transmitter.
- The lost ACK or NAK frames are treated in the same manner as the go-back-n method.
- When the transmitter reaches either the capacity of its window $[(n + 1)/2]$ or the end of its transmission it sets a timer.
- If no acknowledgement arrives in the allotted time, all the frames that remain unacknowledged are retransmitted.
- The disadvantage of this method is that because of the complexity of sorting and storage required by the receiver and the extra logic needed by the transmitter to select frames for retransmission, the system becomes more expensive.
- The advantage of this system is that it gives the best throughput efficiency.
- This is due to the use of pipelining in selective repeat ARQ.

3.13.5 How to improve the throughput efficiency ?

- If the data signalling rate (R) is increased, then the time taken to transmit each block (B/R) will be reduced.
- However as delay remains unchanged, the throughput efficiency will decrease.
- To compensate for this it will be necessary to use longer blocks for higher data rates (R).
- Longer blocks however will have a greater probability of error, therefore an optimum block length is must be obtained for any particular system.
- Throughput efficiency also depends on the type of system used.
- For a half duplex system the transmission efficiency is very poor. An alternative method which gives greater efficiency is to use a continuous mode of transmission instead of block by block transmission.
- In this system the data blocks are transmitted without interruption unless a negative acknowledgement signal (NAK) is received by the transmitting end.
- When NAK is transmitted back to the transmitter it will retransmit the error block. The continuous transmission method avoids the dead time but needs more storage or buffering.

3.13.6 Comparison of Sliding Window Protocols :

MU : May 17

University Questions

Q. 1 With the help of suitable example explain sliding window protocol with selective repeat. Compare its performance to sliding window with Go-back-n technique. **(May 17, 10 Marks)**

Table 3.13.1 : Comparison of sliding window protocols

Sr. No.	Parameter	Stop and wait	Go back n ARQ	Selective repeat ARQ
1.	Window size	1.	Sending window size : $(2^m - 1)$	Sending window size : 2^{m-1}
2.	Operating principle	Transmits one frame at a time and waits for its ACK signal. Transmits the next frame only if ACK is obtained.	It transmits frames continuously till it receives the NAK signal.	Same as Go back n protocol.
3.	Communication type (Direction wise).	Communication is one way (simplex) for the data frames though the ACK frames are allowed to travel in the opposite direction.	Communication is one way (simplex) for the data frames though the NAK frames are allowed to travel in the opposite direction.	Same as Go back n protocol
4.	Retransmission takes place if	1. Received frame is damaged. 2. Transmitted frame is lost 3. ACK is lost	1. Received frame is damaged. 2. Transmitted frame is lost. 3. NAK is lost.	Same as Go back n protocol
5.	Retransmission	Only the damaged or lost frame is retransmitted.	On reception of the NAK signal, the transmitter retransmits all the frames from the one for which the NAK is obtained.	On reception of NAK, only the damaged or lost frame is retransmitted.
6.	Principle of pipelining.	Not used	Used	Used.
7.	Efficiency	Least efficient and slow	Moderately efficient due to pipelining	Most efficient due to pipelining.
8.	Complexity	Less complex	Moderately complex.	Highly complex.

Ex. 3.13.1 : Calculate the throughput for stop-and-wait flow control mechanism if the frame size is 4800 bits, bit rate is 9600 bps and distance between device is 2000 km. Speed of propagation over the transmission is 200,000 km/s.

Soln. :

$$t_f = \frac{\text{Frame size}}{\text{Bit rate}} = \frac{4800}{9600} = 0.5 \text{ sec}$$

$$t_p = \frac{2000}{200000} = 0.01 \text{ sec}$$

We know, $A = t_p / t_f$

$\therefore A = 0.01 / 0.5 = 0.02$

Since, $\eta = 1/(1 + 2A) = 1/(1 + 2 \times 0.02)$

$= 0.96$

$\therefore \% \eta = 96 \%$ ...Ans.

Ex. 3.13.2 : A channel has a bit rate of 4 kbps and propagation delay of 20 msec. For what range of frame sizes does stop and wait gives an efficiency of at least 50 percent ?

Soln. :

Given : Bit rate = 4 kbps,
Propagation delay t_p = 20 mS,
Efficiency $\eta \geq 50 \%$ i.e. $0.5 \leq \eta \leq 1$

To find : Range of frame size.

Step 1 : Calculate value of t_f :

$$\eta = \frac{t_f}{t_f + 2 t_p}$$

For $\eta = 0.5$ we get, $0.5 = \frac{t_f}{t_f + (2 \times 20 \times 10^{-3})}$

$\therefore 0.5\, t_f + 20 \times 10^{-3} = t_f$

$\therefore t_f = 40 \times 10^{-3}$ sec.

- Note that t_f = Transmission time for 1 frame

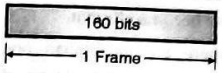
Step 2 : Calculate the frame size :

R = Data rate = 4 kbps = 4000 bps

$\therefore N = R \times t_f$

where N = Frame size

$\therefore N = 40 \times 10^{-3} \times 4 \times 10^3 = 160$ bits ...Ans.



(G-244) **Fig. P. 3.13.2 : 1 frame size**

Ex. 3.13.3 : A channel has a bit rate of 4.8 kbits/sec and a propagation delay of 20 msec. For what range of a frame size does stop and wait protocols given an efficiency of 50%.

Soln. :

Explanation :

- If the channel capacity is B bits/sec, the frame size L bits and the round trip propagation time T seconds, the time required to transmit a single frame is L/B sec.
- After the last bit of a data frame has been sent, there is a delay of at least T/2 for the acknowledgement to come back, for a total delay of T.
- In stop-and-wait the line is busy for T/2 and idle for T, giving an efficiency of L/(L + BT).

Given : Bit rate (B) = 4.8 k bits sec., propagation delay (T) = 20 msec, Efficiency = 50%,

To find : Frame size (L)

$$\text{Efficiency} = \frac{L}{(L + BT)}$$

$$0.5 = \frac{L}{(L + 4.8 \times 10^3 \times 20 \times 10^{-3})} = \frac{L}{(L + 96)}$$

$0.5\,(L + 96) = L$

$0.5L + 48 = L$

$\therefore L = \frac{48}{0.5} = 96$ bits. ...Ans.

Review Questions

- **Q. 1** State the various design issues for the data link layer.
- **Q. 2** State and explain the various services provided to the Network layer.
- **Q. 3** What are the different framing methods ?
- **Q. 4** Explain character stuffing.
- **Q. 5** What is bit stuffing ?
- **Q. 6** Write a note on error control.
- **Q. 7** Explain the simplex protocol for noisy channel.
- **Q. 8** What is piggybacking ?
- **Q. 9** Write a note on sliding window protocols.
- **Q. 10** Explain the stop and wait protocol.
- **Q. 11** State drawbacks of stop and wait protocol.
- **Q. 12** Explain the Go back n protocol.
- **Q. 13** What is pipelining ?
- **Q. 14** Write a note on : Selective repeat ARQ.
- **Q. 15** Define throughput efficiency and explain how it can be increased ?

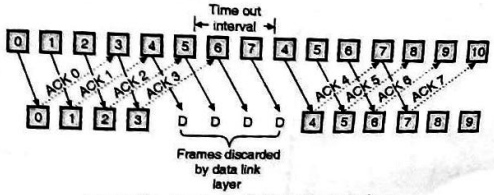
3.14 University Questions and Answers :

- **Q. 1** Explain sliding window protocol. Draw the sender and receiver windows for a system using Go-Back-N sliding window system given that
  1. Frame 0 is sent ; Frame 0 is ACK
  2. Frames 1 and 2 are sent ; Frames 1 and 2 are ACK.
  3. Frames 3, 4, 5 are sent. Frame 4 is ACK. Timer for frame 5 expires.
  4. Frames 5, 6, 7 are sent, Frames 4 through 7 are ACK. **(Dec. 2004, Dec. 2013, 10 Marks)**

Ans. :

- Refer section 3.13 for sliding window protocol.

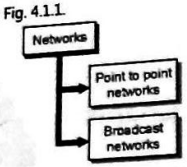
Sender and receiver windows :



(G-803) **Fig. 1 : Go-Back-N sliding window**

4.1 Introduction :

- We can classify the networks into two categories as shown in Fig. 4.1.1.

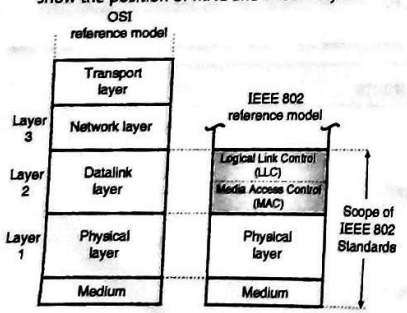


(G-264) Fig. 4.1.1 : Classification of networks

- In this chapter, we are going to discuss the broadcast networks and their protocols.
- The broadcast channels are also called as multi-access channels or random access channels.
- In the broadcast networks the most important point is the criteria by which we decide, who is allowed to use the common channel when more than one user want to use it.
- A protocol is used to make this decision.
- Such a protocol belongs to a sublayer of data link layer called the MAC (Medium Access Control) sublayer.
- The MAC sublayer is very important in LANs because it is a broadcast network.

4.1.1 MAC and LLC Sublayers :

- Fig. 4.1.2 shows the layered OSI model (partial) to show the position of MAC and LLC sublayers.



(G-265) Fig. 4.1.2 : IEEE 802 protocol layers compared to OSI model

- We will discuss the broadcast protocols corresponding to the lower layers (1 and 2) of the OSI model as shown in Fig. 4.1.2.
- Fig. 4.1.2 relates the LAN protocols with the OSI architecture.
- This architecture was developed by IEEE 802 committee and it has been accepted as LAN standard.
- It is called as IEEE 802 reference model. Let discuss this model layer by layer.

Functions of Media Access Control (MAC) sublayer :

- The functions of MAC sublayer are as follows :
  1. To perform the control of access to media.
  2. It performs the unique addressing to stations directly connected to LAN.
  3. Detection of errors.

Functions of Logical Link Control (LLC) sublayer :

- The functions of LLC sublayer are as follows :
  1. Error recovery.
  2. It performs the flow control operation.
  3. User addressing.

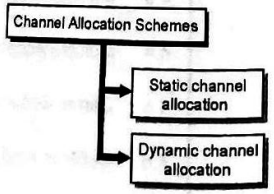
4.2 The Channel Allocation Problem :

Channel allocation :

- In a broadcast network, the single communication channel is to be allocated to only one transmitting user at a time.
- The other users connected to this medium should wait till the transmission medium becomes idle again, otherwise, the transmitted packets from multiple sources would collide with each other and lost.
- This is called as channel allocation.

Channel allocation schemes :

- There are two different schemes used for channel allocation as shown in Fig. 4.2.1.



(G-266) Fig. 4.2.1 : Channel allocation schemes

- The channel allocation schemes are :
  1. Static channel allocation and
  2. Dynamic channel allocation

4.2.1 Static Channel Allocation :

- The traditional way of allocating a single channel, among many users is by means of frequency division multiplexing (FDM).
- The Frequency Division Multiplexing (FDM) and Time Division Multiplexing (TDM) are the examples of static channel allocation.
- In these methods either a fixed frequency band or a fixed time slot is allotted to each user.
- Thus, either the entire available bandwidth or the entire time is shared.

Problems in static allocation :

- Some of the major problems with the static channel allocation schemes are as follows :

1. Wastage of channel bandwidth :

- The problem in the static channel allocation methods is that if all the N users are not using the channel, the channel bandwidth is wasted.

2. Lack of channel bandwidth :

- If there are more than N users who want to use the channel, they cannot do so for the lack of bandwidth.

3. Low efficiency :

- For a small number or users and light traffic the static FDM is an efficient method of allocation but its performance is poor for large number of users.

4. Poor performance for bursty traffic :

- The static channel allocation has a poor performance with bursty traffic.
- Therefore, generally dynamic channel allocation is used, for computer networks where the traffic is of bursty nature.

Performance of Static Allocation Schemes :

- To see the poor performance of static channel, let us consider an example of a FDM system.
- Let the mean time delay be (T) for a channel of capacity C bps, with an arrival rate of λ frames/sec.
- Each frame having a length drawn from an exponential probability density function with mean $1/\mu$ bits/frame is given as,

$$T = \frac{1}{\mu C - \lambda}$$

- If the single channel is divided into N independent sub-channels the above equation is modified as follows :

$$T_{FDM} = \frac{1}{\mu(C/N) - (\lambda/N)}$$

$$= \frac{N}{\mu C - \lambda}$$

$$T_{FDM} = NT$$

- From the above equation, it is clear that the mean delay using FDM becomes worse with increase in the number of users N.
- The dynamic channel allocation is used, for computer networks where the traffic is of bursty nature.

4.2.2 Dynamic Channel Allocation :

- In this method either a fixed frequency or fixed time slot is not allotted to the user.
- The user can use the single channel as per his requirement.

Assumptions :

- Following assumptions are made for the implementation of dynamic channel allocation method :

1. Station model :

- This model consists of N independent stations such as a PC, computer etc. which can generate frames for transmission.

2. Single channel :

- A single channel is available for all communication.

3. Collision :

- If frames are transmitted at the same time by two or more stations, there is an overlap in time and the resulting signal is garbled. This is called as collision.

4. **Continuous or slotted time :**
   - There is no master clock used to divide time into discrete time intervals. So frames can begin at any random instant.
   - This is continuous time. For a slotted time, the time is divided into discrete time slots.
5. **Carrier or No carrier sense :**
   - Stations sense the channel before transmission or they directly transmit without sensing the channel.

4.3 Multiple Access :

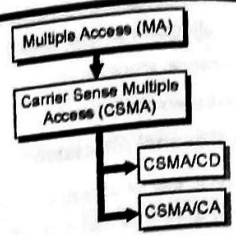
- When a number of stations (users) use a common link of communication system we have to use a multiple access protocol in order to coordinate the access to the common link.

Types of Multiple Access Techniques :

- The three techniques used to deal with the multiple access problem are as follows :
  1. Random Access
  2. Controlled Access
  3. Channelization

4.3.1 Random Access :

- In the random access technique there is no control station.
- Each station will have the right to use the common medium without any control over it.
- With increase in number of stations, there is an increased probability of **collision** or **access conflict.**
- The collisions will occur when more than one user tries to access the common medium simultaneously.
- As a result of such collisions some frames can be either modified (due to errors) or destroyed.
- In order to avoid collisions, we have to set up a procedure.
- The evolution of the random access methods is shown in Fig. 4.3.1.



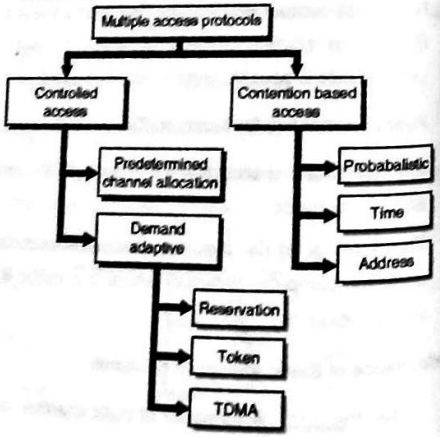
(G-267) **Fig. 4.3.1 : Evolution of random access methods**

4.3.2 Evolution of Random Access Methods :

- The first method in the evolution ladder of Fig. 4.3.1, known as ALOHA used a simple procedure called multiple access (MA).
- It was improved to develop the carrier sense multiple access (CSMA).
- The CSMA further evolved into two methods namely CSMA/CD (CSMA with collision detection) and CSMA/CA (CSMA with collision avoidance) which avoids the collisions.

4.3.3 Taxonomy (Classification) of Multiple Access Protocols :

- Multiple access protocols, determine the allocation of channel among the stations. These protocols can be broadly divided into two classes :
  1. Controlled access
  2. Contention based access.
- Fig. 4.3.2 explains these two classes and their further sub-division.



(G-1544) **Fig. 4.3.2 : Classification of Multiple Access Protocols**

- The controlled access protocols are characterized by a collision free access to the channel.
- That means the stations are co-ordinated in such a way that two or more stations never attempt to transmit simultaneously.
- The controlled access protocols are further classified into two types :
  1. Demand adaptive
  2. Predetermined channel allocation.
- The PCA (TDMA) protocols allocate the channel to the stations in a static manner (on fixed time sharing basis).
- The demand adaptive protocols try to allocate the channel as per the demands of the stations.
- The controlled access protocols have been discussed in section 4.7.
- Both PCA and demand adaptive protocols are inefficient if there are large number of stations having bursty message transmissions.
- The second broad class of multiple access protocols is known as contention-based protocols.
- These are further classified into three types, probabilistic, them based and address based protocols.
- The ALOHA and CSMA protocols (Sections 4.4 and 4.5) belong to the contention based protocols.

4.4 Multiple Access ALOHA System :

MU : Dec. 06, May 10, Dec. 10, May 11, May 13

University Questions

Q. 1 Write short notes on : ALOHA. **(Dec. 06, 3 Marks)**

Q. 2 Explain ALOHA and Slotted ALOHA in detail. **(May 10, 10 Marks)**

Q. 3 Explain ALOHA in detail. **(Dec. 10, May 11, 5 Marks)**

Q. 4 Explain the ALOHA protocol. Compare the performance of pure ALOHA versus slotted ALOHA at low load and high load. **(May 13, 10 Marks)**

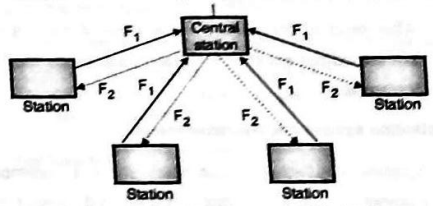
ALOHA System :

- Systems in which multiple users share a common channel in a way that can lead to conflicts are widely known as **contention systems.**
- The ALOHA system is a contention protocol which was developed at the University of Hawaii in the early 1970's by Norman Abramson and his colleagues.
- The ALOHA system has two versions :
  1. Pure ALOHA – Does not require global time synchronisation.
  2. Slotted ALOHA – Requires time synchronisation.

4.4.1 Pure ALOHA :

Principle :

- It works on a very simple principle. Essentially it allows for any station to broadcast at any time.
- If two signals collide, each station simply waits a random time and try again.
- Collisions are easily detected. As shown in the Fig. 4.4.1, when the central station receives a frame it sends an acknowledgement on a different frequency.



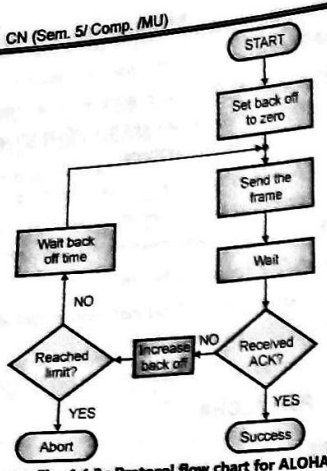
F_1 = Broadcast frequency from the individual stations.
F_2 = Broadcast frequency from the central station.

(G-268) **Fig. 4.4.1 : Pure ALOHA system**

- If a user station receives an acknowledgement it assumes that the transmitted frame was successfully received and if it does get an acknowledgement it assumes that collision had occurred and is ready to retransmit.
- The advantage of pure ALOHA is its simplicity in implementation but it's performance becomes worse as the data traffic on the channel increases.

Protocol Flow Chart for ALOHA :

- Fig. 4.4.2 shows the protocol flow chart for ALOHA.



(G-269)Fig. 4.4.2 : Protocol flow chart for ALOHA

Explanation :

- A station which has a frame ready for transmission will send it and waits for some time.
- If it receives an acknowledgement then the transmission is successful.
- Otherwise the station uses a **back-off** strategy, and will send the packet again.
- After sending the packet, many times if there is no acknowledgement then the station aborts the idea of transmission.

Contention system and Retransmission :

- Systems in which multiple users share a common channel in such a way that can lead to a conflict or collision are known as the contention systems.
- Whenever two frames try to occupy the channel at the same time, there is bound to be a collision and both will be garbled.
- Retransmission is essential for all the destroyed frames.

4.4.2 Efficiency of an ALOHA System :

MU : Dec. 04, May 12

University Questions

Q. 1 Derive the formula for measuring the efficiency of the ALOHA system and explain how the efficiency is increased for slotted ALOHA. **(Dec. 04, 10 Marks)**

Q. 2 Derive the efficiency of Pure ALOHA protocol. **(May 12, 10 Marks)**

- Efficiency of an ALOHA system is that fraction of all transmitted frames which escape collisions i.e. which do not get caught in collisions.
- Consider ∞ number of interactive users at their computers (stations).
- Each user is either typing or waiting. Initially all of them are in the typing state.
- When a user types a line, the user stops and waits. The station then transmits a frame containing this line and checks the channel to confirm the success.
- If it is successful then the user will start typing again, otherwise the user waits and its frame is retransmitted many time till it is sent successfully.

Frame time :

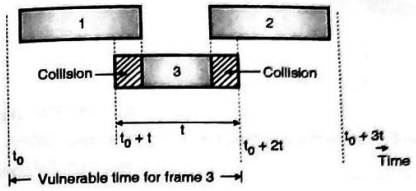
- Let the frame time be defined as the amount of time required to transmit the standard fixed length frame. Note that

$$\text{Frame time} = \frac{\text{Frame length}}{\text{Bit rate}}$$

- We assume that ∞ number of users generate new frames according to the Poisson's distribution with an average N frames per frame time.
- The value of N > 1 indicates that the users are generating frames at a rate higher than that can be handled by the channel.
- So most of the frames will face collision. Hence 0 < N < 1 in order to reduce number of collisions.
- Let there be k transmission attempts (including retransmissions) per frame time.
- The probability of k transmissions per frame time is also Poisson. Let the mean of number of transmissions be G per frame time. So G ≥ N.
- At low load N ≈ 0 there will be less number of collisions so less number of retransmissions and G ≈ N.
- With increase in load there are many collisions so G > N. Combining all these we can say that for all the loads the throughput is given by,

$$S = GP_0$$

- Where P_0 = Probability that a frame does not suffer a collision. Consider Fig. 4.4.3.



(G-270) Fig. 4.4.3

- What is the condition for frame 3 in Fig. 4.4.3 to arrive undamaged without collision ?
- Let t = Time required to send a frame. If frame 1 is generated at any instant between t_0 to $(t_0 + t)$ then it will collide with frame 3. Similarly any frame (2) generated between $(t_0 + t)$ and $(t_0 + 2t)$ also collides with frame 3.
- As per Poisson's distribution, the probability of generating k frames during a given frame time is given by,

$$P[k] = \frac{G^k e^{-G}}{k!}$$

- So the probability of generating zero frames i.e. k = 0 is

$$P_0 = \frac{G^0 e^{-G}}{0!} = e^{-G}$$

- If an interval is two frame time long, the mean number of frames generated during that interval is 2G.
- The probability that no other frame is transmitted during the Vulnerable period (time when collision can take place) is,

$$P_0 = e^{-2G}$$

- But throughput $S = G P_0$

$$\therefore \quad S = Ge^{-2G}$$

- Fig. 4.4.5 shows the relation between the offered traffic G and the throughput S.
- It shows that the maximum throughput occurs at G = 0.5 and S_{max} = 0.184. So the best possible channel utilization is on 18.4 percent.

4.4.3 Slotted ALOHA :

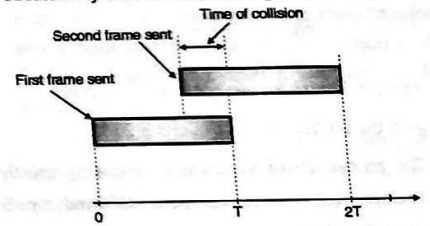
MU : Dec. 07, May 10

University Questions

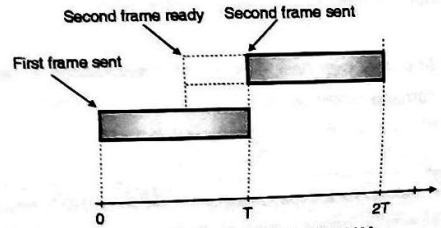
Q. 1 Describe in brief : Slotted ALOHA. **(Dec. 07, 5 Marks)**

Q. 2 Explain ALOHA and Slotted ALOHA in detail. **(May 10, 10 Marks)**

- To overcome the disadvantage of the pure ALOHA system (of low capacity) Robert published a method for doubling the capacity of traffic on the channel.
- In this method it was proposed that the time be divided up into discrete intervals and each interval correspond to one frame.
- This method requires that the users agree on the slot boundaries.
- In this method for achieving synchronisation one special station emits a pip at the start of each interval, like a clock.
- This method is known as the slotted ALOHA system.
- Collisions occur if any part of two transmission overlaps.
- Suppose that T is time required for one transmission and that two stations must transmit.
- The total time required for both stations to do so successfully is 2T as shown in Fig. 4.4.4.



(a) Transmission using pure ALOHA



(b) Transmission using slotted ALOHA

(G-271) Fig. 4.4.4

- In case of pure ALOHA allowing a station to transmit at arbitrary times can waste time upto 2T.
- As an alternative, in the slotted ALOHA method the time is divided into intervals (slots) of T units each and require each station to begin each transmission at the beginning of a slot.
- In other words, even if station is ready to send in the middle of a slot, it must wait until the beginning of the next one as shown in Fig. 4.4.4(b).
- In this method a collision occurs when both stations become ready in the same slot.
- Slotted ALOHA is thus a discrete time system whereas pure ALOHA is a continuous time system.
- The Vulnerable period has been reduced to half that of pure ALOHA, the throughput for slotted ALOHA is given by,

$$S = Ge^{-G}$$

- The maximum throughput corresponds to G = 1 and it is given by $S_{max} = 1/e = 0.368$ as shown in Fig. 4.4.5.
- So for a slotted ALOHA with G = 1 the probability of success is 37%.
- The probability of empty slots is,

$$P(k) = \frac{G^k e^{-G}}{k!}$$

For G = 1 and k = 0 we get P (k = 0) = 0.368.

And the probability of collisions is 26 %.

- The probability of transmission requiring exactly k attempts (i.e. k – 1 collisions followed by one success) is given by,

$$P_k = e^{-G}(1-e^{-G})^{k-1}$$

- And the expected number of transmissions E per carriage return typed is

$$E = e^{G}$$

Conclusion : As E depends exponentially on G, with a small increase in G, there is a large increase in E and drastic fall in performance.

4.4.4 Comparison of Pure and Slotted ALOHA :

MU : May 13, May 15

University Questions

Q. 1 Explain the ALOHA protocol. Compare the performance of pure ALOHA versus slotted ALOHA at low load and high load. **(May 13, 10 Marks)**

Q. 2 Differentiate between ALOHA and slotted ALOHA. **(May 15, 4 Marks)**

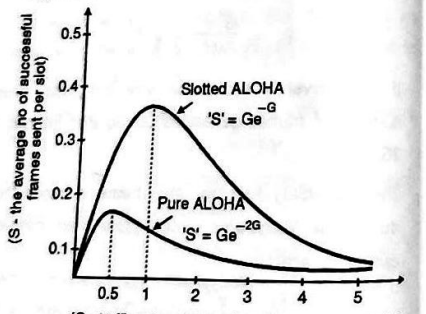
- A mathematical model can be created for the relationship between the number of frames transmitted and the number of frames transmitted successfully.
- Let G represent the traffic measured as the average number of frames generated per slot.
- Let S be the success rate measured as the average number of frames sent successfully per slot.
- The relationship between G and S for both pure and slotted ALOHA is given as follows :

Pure ALOHA → $S = Ge^{-2G}$

Slotted ALOHA → $S = Ge^{-G}$

Where e is the mathematical constant = 2.718.

- From the above equation a success rate curve for pure and slotted ALOHA can be plotted as shown in Fig. 4.4.5.



(G-272)Fig. 4.4.5 : Comparison of pure and slotted ALOHA

- As seen in the Fig. 4.4.5 both graphs have the same shape.
- If G is small so is S, which means that if few frames are generated few frames will be transmitted successfully.

- As G increases so does S but upto a certain point. As G continues to increase S approaches to 0 which means that if more frames are generated there will be more collisions and the success rate will fall to 0.
- Similarly for pure ALOHA the maximum occurs at G = 0.5 for which S = 1/2e = 0.184 which means the rate of successful transmissions is approximately 18.4%.
- As seen from the graph the maximum for slotted ALOHA occurs at G = 1 for which S = 1/e = 0.368.
- In other words the rate of successful transmissions is approximately 0.368 frames per slot time or 37% of the time will be spent on successful transmissions.
- Hence the slotted ALOHA has a double throughput efficiency than the pure ALOHA system.
- The maximum utilization achievable using CSMA can be increased much beyond that obtainable using ALOHA or slotted ALOHA.
- The maximum utilization is dependent on length of the frame and on the propagation time.
- With increase in the length of the frame or reduction in the propagation time the utilization gets improved.

Ex. 4.4.1 : A pure ALOHA network transmits 200 bit frames on a shared channel of 200 kbps.

What is the throughput if the system (all stations together) produces ?

1. 1000 frames per second
2. 500 frames per second
3. 250 frames per second. **(Dec. 16, 10 Marks)**

Soln. :

Given : Rate of transmission = 200 kbps = 200000 bps

Frame length = 200 bits.

To find : Throughput

1. Number of frames / sec = 1000 frames / sec.
   - The maximum throughput for a pure ALOHA system is 0.184

   ∴ Throughput = 1000 × 0.184 = 184 frames /sec
2. Number of frames / sec = 500

   ∴ Throughput = 500 × 0.184 = 92 frames / sec
3. Number of frames / sec = 250

   ∴ Throughput = 200 × 0.184 = 36.8 frames / sec

Throughput of all stations = 184 + 92 + 36.8

= 312.8 frames / sec

4.5 Carrier Sense Multiple Access (CSMA) :

MU : Dec. 04, Dec. 13, Dec. 14, May 15, Dec. 15, Dec. 17, Dec. 18, Dec. 19

University Questions

Q. 1 How are collisions handled by a 1-persistant CSMA protocol ? Give an example of a collision-free multiple access protocol and explain it in detail. **(Dec. 04, 10 Marks)**

Q. 2 Explain the different protocols in the MAC sublayer which uses carrier sensing. **(Dec. 13, 10 Marks)**

Q. 3 Explain CSMA protocols. Explain how collisions are handled in CSMA/CD. **(Dec. 14, May 15, Dec. 16, Dec. 17, 10 Marks)**

Q. 4 Explain CSMA protocols. Explain how collision are handled in CSMA / CD. **(Dec. 18, Dec. 19, 10 Marks)**

- The CSMA protocol operates on the principle of carrier sensing.
- In this protocol, a station listens to see the presence of transmission (carrier) on the cable and decides to act accordingly.

Non-Persistent CSMA :

- In this scheme, if a station wants to transmit a frame and it finds that the channel is busy (some other station is transmitting) then it will wait for fixed interval of time.
- After this time, it again checks the status of the channel and if the channel is free it will transmit.

1-Persistent CSMA :

- In this scheme the station which wants to transmit, continuously monitors the channel until it is idle and then transmits immediately.
- The disadvantage of this strategy is that if two stations are waiting then they will transmit simultaneously and collision will take place. This will then require retransmission.

P-Persistent CSMA :

- The possibility of such collisions and retransmissions is reduced in the p-persistent CSMA.

- In this scheme all the waiting stations are not allowed to transmit simultaneously as soon as the channel becomes idle.
- A station is assumed to be transmitting with a probability "p".
- For example if p = 1/6 and if 6 stations are waiting then on an average only one station will transmit and others will wait.

4.5.1 Carrier Sense Multiple Access/Collision Detection (CSMA/CD) :

MU Dec. 13, Dec. 14, May 15, Dec. 15, Dec. 16, Dec. 17, Dec. 18, Dec. 19

University Questions

Q. 1 Explain the different protocols in the MAC sublayer which uses carrier sensing. **(Dec. 13, 10 Marks)**

Q. 2 Explain CSMA protocols. Explain how collisions are handled in CSMA/CD. **(Dec. 14, May 15, Dec. 16, Dec. 17, 10 Marks)**

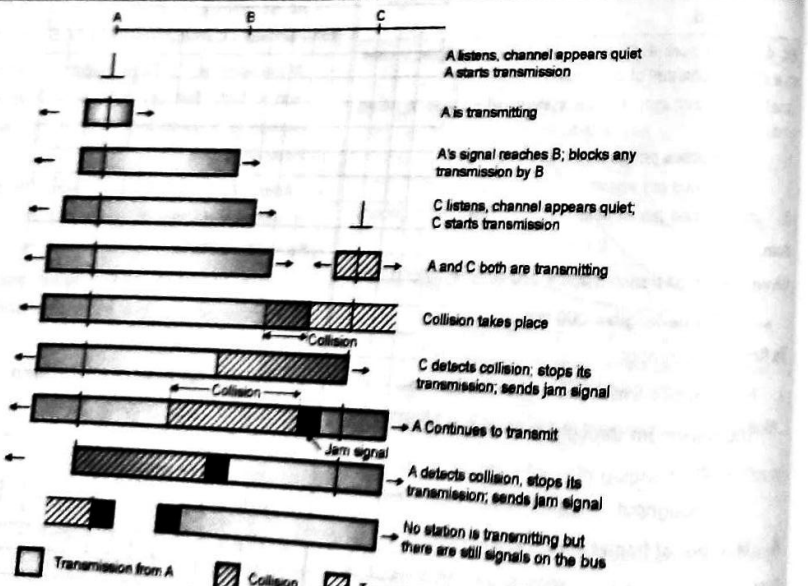
Q. 3 Write in brief about : CSMA/CD **(Dec. 15, 5 Marks)**

Q. 4 Explain CSMA protocols. Explain how collision are handled in CSMA / CD. **(Dec. 18, Dec. 19, 10 Marks)**

- The CSMA/CD specifications have been standardized by IEEE 802.3 standard.
- It is a very widely used MAC protocol.

Media access control :

- The problem in CSMA explained earlier is that a transmitting station continues to transmit its frame even though a collision occurs.
- The channel time is unnecessarily wasted due to this. In CSMA/CD, if a station receives other transmissions when it is transmitting, then a collision can be detected as soon as it occurs and the transmission time can be saved.
- As soon as a collision is detected, the transmitting stations release a jam signal.
- The jam signal will alert the other stations.
- The stations then are not supposed to transmit immediately after the collision has occurred.
- Otherwise there is a possibility that the same frames would collide again.
- After some "back off" delay time the stations will retry the transmission.
- If again the collision takes place then the back off time is increased progressively.
- A careful design can achieve efficiencies of more than 90% using CSMA/CD.
- This scheme is as shown in Fig. 4.5.1.



(G-273)Fig. 4.5.1 : CSMA/CD scheme

4.5.2 CSMA/CD Procedure :

MU Dec. 14, May 15, Dec. 15, Dec. 16, Dec. 17, Dec. 18, Dec. 19

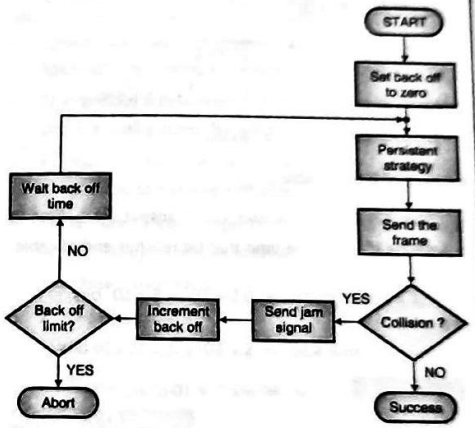
University Questions

Q. 1 Explain CSMA protocols. Explain how collisions are handled in CSMA/CD. **(Dec. 14, May 15, Dec. 16, Dec. 17, 10 Marks)**

Q. 2 Write in brief about : CSMA/CD **(Dec. 15, 5 Marks)**

Q. 3 Explain CSMA protocols. Explain how collision are handled in CSMA / CD. **(Dec. 18, Dec. 19, 10 Marks)**

- Fig. 4.5.2 shows a flow chart for the CSMA/CD protocol.



(G-276)Fig. 4.5.2 : CSMA/CD procedure

Explanation :

- The station that has a ready frame sets the back off parameter to zero.
- Then it senses the line using one of the persistent strategies.
- It then sends the frame, if there is no collision for a period corresponding to one complete frame, then the transmission is successful.
- Otherwise (in the event of collision) the station sends the jam signal to inform the other stations about the collision.
- The station then increments the back off time and waits for a random back off time and sends the frame again.
- If the back off has reached its limit then the station aborts the transmission.
- CSMA/CD is used for the traditional Ethernet.
- CSMA/CD is an important protocol. IEEE 802.3 (Ethernet) is an example of CSMA/CD. It is an international standard.
- The MAC sublayer protocol does not guarantee reliable delivery.
- Even in absence of collision the receiver may not have copied the frame correctly.

4.5.3 CSMA/CA :

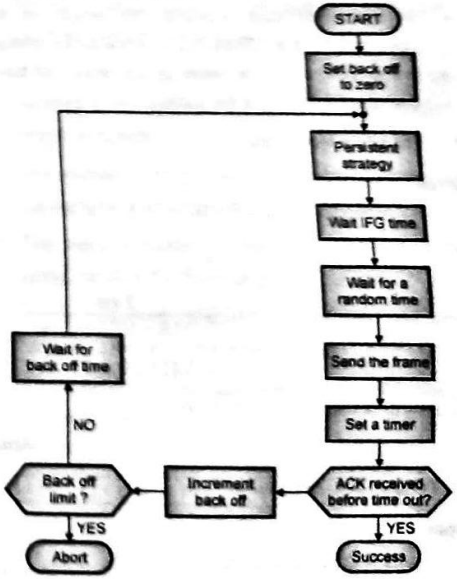
MU Dec. 09, Dec. 13

University Questions

Q. 1 What is Carrier Sense Multiple Access with Collision Avoidance (CSMA/CA) protocol ? Explain with timing diagram. **(Dec. 09, 10 Marks)**

Q. 2 Explain the different protocols in the MAC sublayer which uses carrier sensing. **(Dec. 13, 10 Marks)**

- The long form of CSMA/CA is CSMA protocol with collision avoidance.
- Fig. 4.5.3 shows the flow chart explaining the principle of CSMA/CA.



(G-277)Fig. 4.5.3 : CSMA/CA procedure

- The station ready to transmit, senses the line by using one of the persistent strategies.
- As soon as it find the line to be idle, the station waits for a time equal to an IFG (Interframe gap).
- It then waits for some more random time and sends the frame.
- After sending the frame, it sets a timer and waits for the acknowledgement from the receiver.
- If the acknowledgement is received before expiry of the timer, then the transmission is successful.
- But if the transmitting station does not receive the expected acknowledgement before the timer expiry then it increments the back off parameter, waits for the back off time and senses the line again. CSMA/CA completely avoids the collision.

Ex. 4.6.1 : What is the length of a contention slot in CSMA / CD for : 1. 2-km twin-lead cable (signal propagation speed is 82% of the signal propagation speed in vacuum) and 2. A 40-km multimode fiber optic cable (signal propagation speed is 65% of the signal propagation speed in vacuum).

Soln. :

- In the worst case scenario, the length of the contention slot in CSMA / CD is equal to 2τ, where τ is the minimum time taken by the signal to travel over the full length of the cable.

1. For a twin lead cable :

Given : L = 2 km,

Speed $S = 0.82 \times 3 \times 10^8 = 24.6 \times 10^5$ km /s

$$\therefore \tau = \frac{L}{S}$$

$$= \frac{2 \text{ km}}{24.6 \times 10^5 \text{ km / sec}}$$

$$= 0.813 \ \mu S.$$

$\therefore$ Length of contention slot $= 2\tau$

$= 1.626 \ \mu S$ **...Ans.**

2. For the fiber optic cable :

Given : L = 40 km,

$S = 0.65 \times 3 \times 10^8$ m/S

$= 1.95 \times 10^5$ km/s.

$$\therefore \tau = \frac{L}{S} = \frac{40}{1.95 \times 10^5}$$

$= 2.05 \times 10^{-4}$ sec.

$\therefore$ Length of contention slot $= 2\tau$

$= 4.1 \times 10^{-4}$ sec

Ex. 4.6.2 : 1 Gbps CSMA/CD LAN is to be designed over 1 km cable without repeater. The cable supports signal speed of 200,000 km/sec. What is the minimum frame size that Data Link Layer should consider.

Soln. :

Propagation speed = 200000 km/sec.

Length of cable = 1 km

Propagation Time $= \frac{1}{200000}$

$= 5 \times 10^{-6}$ s = 5 μ sec

Transmission speed = 1 Gbps.

Number of bits in cable :

- Number of bits sender can transmit from time it sends 1st bit to the time that bit reaches end of cable.

$1 \times 10^9 \times \frac{1}{20000} = 0.5 \times 10^5 = 5 \times 10^4$ bits.

Frame size $= 5 \times 10^4 \times 2 = 10{,}000$ bits

Total round time $= 5 \times 2 = 10$ μ sec.

- For collision detection frame should take at least 10 μS to send.

Data rate = 1000 bit per μS.

- Thus 10,000 bits could be sent in 10 μS. Thus frame size should be at least 10,000 bits.

4.6 Collision Free Protocols :

- As we have seen that almost collisions can be avoided in CSMA/CD, they can still occur during the contention period.
- The collision during contention period affects the systems performance adversely.
- This happens when the cable is long and length of frames is short.

- This problem becomes serious as fiber optic networks come into use.
- Some protocols that resolve the collisions during the contention period are as follows :
  1. Bit map protocol
  2. Binary Countdown
  3. Limited Contention Protocols

4.7 Controlled Access : MU : Dec. 15

University Questions

Q. 1 What is controlled access for collision control ? Explain all the methods of controlled access. (Dec. 15, 10 Marks)

- Earlier we have discussed the random access approach for sharing a transmission medium.
- The random access approach is simpler to implement and are useful in handling the light traffic.
- In this section we will discuss the scheduling approaches to the medium access control.
- There are three important approaches in the scheduling approach as follows :
  1. Reservation system
  2. Polling system
  3. Token passing ring networks.

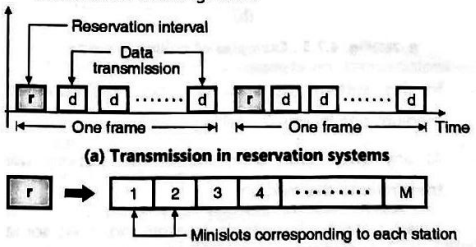
4.7.1 Reservation Systems : MU : Dec. 15

University Questions

Q. 1 What is controlled access for collision control ? Explain all the methods of controlled access. (Dec. 15, 10 Marks)

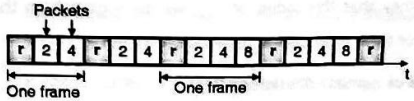
Principle :

- The principle of reservation system can be understood from Fig. 4.7.1.



(L-733)Fig. 4.7.1 : Basic reservation system

- In this system each station transmits a single packet at the full rate R bps.
- The transmissions from the stations can be organized into frames of variable length.
- Before each frame a reserved slot or reservation interval is transmitted as shown in Fig. 4.7.1(a).
- Fig. 4.7.1(b) shows the details of the reservation interval "r".
- The reservation interval consists of M minislots with one slot allotted to each station.
- These minislots are used by the stations to indicate that they have a packet to transmit in the corresponding frame.
- The station that wants to transmit packet by broadcasting their reservation bit during the appropriate minislot.
- All the stations will listen to the reservation interval, and then determine the order in which packet transmissions in the corresponding frame would take place.
- The frame length would correspond to the number of stations which have a packet to transmit.
- If the length of the packet is variable, then it can be handled if the reservation message includes packet length information.
- This reservation system that we discussed is called as the basic reservation system.
- The basic reservation system can be improved by using the time division multiplexing scheme.
- In the improved reservation system the idle time slots are allotted to the other stations.
- The operation of the basic reservation system can be explained with the help of Fig. 4.7.2.



(a) Negligible propagation delay

(L-734)Fig. 4.7.2 (Contd..)

(b) Non negligible propagation delay

(L-734)Fig. 4.7.2 : Operation of reservation system with negligible and non-negligible delays

- Refer Fig. 4.7.2(a) which shows a system with negligible propagation delay.
- In the first frame, only the stations 2 and 4 transmit their packets.
- But in the middle portion, station 8 also wants to transmit its packet. So the frame gets expanded from two slots to three slots.
- The maximum throughput from this system can be attained when all the stations transmit their packet in each frame.
- The corresponding maximum throughput is given by,

$$\rho_{max} = \frac{1}{1+v}$$ ...for one packet reservation/minislot

- If $v << 1$ then the value of ρ_{max} can be very high.
- Now refer Fig. 4.7.2(b) which shows a reservation system with some finite non zero propagation delay which can not be neglected.
- In this system the stations will transmit their reservations in the same way as they used to do before.
- It is possible to modify the basic reservation system so that stations can reserve more than one slot per packet transmission per minislot.
- Let us assume that a minislot can reserve say upto k packets.
- Then the maximum achievable throughput is given by,

$$\rho_{max} = \frac{1}{1+(v/k)}$$ ...for k packet reservation/minislot

- Note that this value of ρ_{max} will be higher than that for the single packet reservation/minislot.

Effect of number of stations (M) :

- The reservation intervals introduce overhead which is proportional to M.
- That means the reservation interval becomes M × v.
- As the number of stations (M) become very large, this overhead will become significant.
- This then becomes a serious problem.
- This problem can be sorted out by not allocating a minislot to each station and then instead making the stations to compete for a reservation of minislot by using a random access technique such as ALOHA or slotted ALOHA.

4.7.2 Polling :

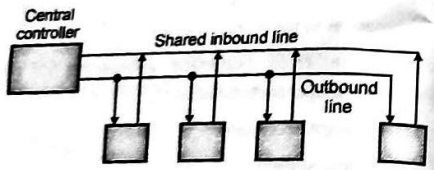
MU : Dec. 15

University Questions

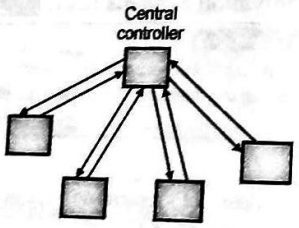
Q. 1 What is controlled access for collision control ? Explain all the methods of controlled access. **(Dec. 15, 10 Marks)**

Principle :

- Now consider polling system shown in Fig. 4.7.3.



(a)

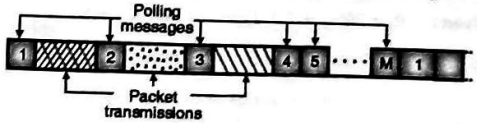


(b)

(L-735)Fig. 4.7.3 : Examples of polling systems

- In this system the stations access the common medium one by one (by taking turns).
- At any given time only one of the stations will transmit into the medium.
- When a station finishes its transmitting, then some mechanism is used to pass the right of transmission to another station which wants to transmit next.

- There are different ways of passing the right of transmission from one station to the other station.
- Fig. 4.7.3(a) shows a scheme in which M stations communicate with a central controller.
- The outbound line is used for carrying the information from the central controller to the M users whereas the shared inbound line is required to carry the information from users to the central computer.
- Thus the inbound line acts as the shared medium that requires a medium access control (MAC).
- The host computer acts as a central controller.
- It sends control messages which co-ordinate the transmissions from the stations.
- The central controller sends a polling message to a particular station.
- That station sends its message on the shared inbound line. Once this process is over, the station gives a go-ahead message.
- It is possible that the central controller may poll the stations in a round robin (serial) fashion or it may do it according to some pre-determined rule.
- Fig. 4.7.3(b) shows another system where it is possible to use polling.
- The central controller of this system can make use of radio transmission.
- Fig. 4.7.4 shows the sequence of polling messages.



(L-736)Fig. 4.7.4 : Polling messages and transmissions in a polling system

- Station 1 gets the polling message first. The polling message will propagate.
- It is received by all stations but only station 1 begins transmission. All this process needs a time called **walk time.**
- The next period is occupied by the transmission from station 1.
- This period will then be followed by the walk time corresponding to station-2.
- This process will continue until all the M stations are polled.
- Thus in this system the stations are polled in the round robin manner.
- The walk time can be considered to be an overhead in the polling system because it is an unproductive time.
- The total walk time 't' is the sum of walk time corresponding to each station.

4.7.3 Token Passing :

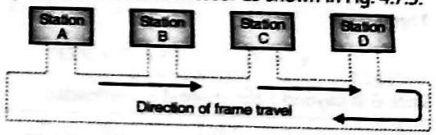
MU : Dec. 15

University Questions

Q. 1 What is controlled access for collision control ? Explain all the methods of controlled access. **(Dec. 15, 10 Marks)**

Principle :

- **Token** is a special frame which is used to authorize a particular station for transmission.
- In the token passing method, the token is given to that station, which is authorized to send its data.
- Thus the station that has the token with it can transmit others listen.
- In a token passing network, each station has a predecessor and successor as shown in Fig. 4.7.5.



(L-737)Fig. 4.7.5 : Token passing network

- The frames travel in one direction. They come from the predecessor and go to the successor as shown in Fig. 4.7.5.
- A token frame is circulated around the ring when no data is being transmitted and the line is idle.
- The stations which are ready to send data, will wait for the token.

- As the token circulates the first ready station in the ring will grab the circulating token and transmit one or more frames.
- This station will keep sending the frames as long as it has frames to send or the allotted time is not complete.
- It then passes this token on the ring from which the next ready to transmit station will grab it.
- This is the simplest possible token passing technique in which all the stations have equal priority or right to send.
- In the practical system, some other features such as priority and reservation are added.

Ex. 4.7.1 : Measurements of a slotted ALOHA channel with an infinite number of user. Show that 10% of the slots are idle.

1. What is channel load ?
2. What is throughput ?
3. Is the channel overloaded or underloaded.

Soln. :

1. Channel load :

For a slotted ALOHA, $P_0 = e^{-G}$

But $P_0 = 10\%$ i.e. 0.1

$\therefore \quad 0.1 = e^{-G}$

$\therefore \quad -2.3 = -G$

$\therefore G = 2.3$

2. Throughput :

$$S = Ge^{-G} = 2.3\,e^{-2.3} = 0.23$$

3. Since G is beyond 1 the channel is overloaded.

Ex. 4.7.2 : Consider building a CSMA/CD network running at 1Gbps over a 1 km cable with no repeaters. The signal speed in the cable is 2,00,000 km/sec, what is the minimum frame size ?

Soln. :

Given : Bit rate $R = 1 \times 10^9$ Bits/sec.

No repeaters used Length $L = 1$ km $= 1 \times 10^3$ m

Speed $v = 2,00,000$ km/S $= 2 \times 10^8$ m/S

To find : Minimum frame size

1. Let the time for a signal to propagate between two farthest stations be τ. The contention interval is such that width of each slot is 2τ.
2. On a 1 km long cable $\tau \approx 5$ µsec. $\therefore 2\tau = 10$ µsec.
3. To make CSMA/CD work, it must be ensured that the minimum frame size should be equal to $2\tau = 10$ µsec.

But $R = 1 \times 10^9$ bits/sec

$\therefore$ 1 sec $\equiv 1 \times 10^9$ bits

$\therefore \quad 10 \times 10^{-6}$ sec $= ?$ bits

$$\therefore \quad \frac{1}{10 \times 10^{-6}} = \frac{1 \times 10^9}{x}$$

$$\therefore \quad x = 1 \times 10^9 \times 10 \times 10^{-6}$$

$$= 10 \times 10^3$$

$$= 10{,}000 \text{ bits.}$$

$\therefore$ Minimum frame size = 10,000 bits or 1250 bytes.

Ex. 4.7.3 : Measurement of slotted ALOHA channel with an infinite number of users show that 20% slots are idle.

1. What is the channel load ?
2. What is the throughput ?
3. Is the channel underload or overload ? Show with graph.

Soln. :

Given : $P_0 = 20\%$ i.e. 0.2, Type : slotted ALOHA

To find : 1. Channel load G 2. Throughput S.

3. Decide the status of the channel

1. Channel load (G) :

For the slotted ALOHA, $P_0 = e^{-G}$

$\therefore \quad 0.2 = e^{-G}$

$\therefore \quad G = 1.6094$ ...Ans.

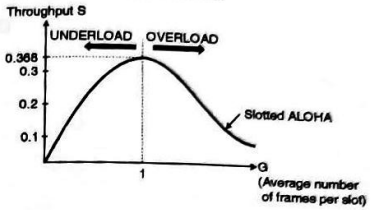
2. Throughput (S) :

$$S = Ge^{-G} = P_0 G = 0.2 \times 1.6094$$

$\therefore \quad S = 0.3218$ ...Ans.

3. Status of the channel :

- From Fig. P. 4.7.3 it is evident that the maximum throughput $S_{max} = 0.368$ corresponds to G = 1.
- Since the value of G = **1.6094** which **is greater** than 1, the channel is **overloaded.**



(L-746)**Fig. P. 4.7.3 : Graph for slotted ALOHA**

Ex. 4.7.4 : ALOHA protocol is used to share 56 kbps satellite channel. If each packet is 1000 bits long find maximum throughput in pockets/sec.

Soln. :

Given : Rate of transmission = 56 kbps = 56000 bps

Frame length = 1000 bits

1. For pure ALOHA :

$$\therefore \text{ Number of frames/sec} = \frac{56000 \text{ bits}}{1000 \text{ bits/frame}}$$

$$= 56 \text{ frames/sec}$$

- The maximum throughput for pure ALOHA = 0.184

$\therefore$ Throughput $= 56 \times 0.184$

$= 10.304$ frames/sec. ...Ans.

2. For slotted ALOHA :

Maximum throughput = 0.368

$\therefore$ Throughput $= 0.368 \times 56$

$= 20.608$ frames/sec. ...Ans.

Ex. 4.7.5 : A group of N users share 56 kbps pure ALOHA channel. Each station outputs 1000 bits frame on an average of once 100 sec. Even if the previous has not yet been sent (buffered) what is maximum value of N ?

Soln. :

For pure ALOHA :

- The maximum throughput = 0.184
- $\therefore$ The maximum usable channel bandwidth is given by,

$R = 0.184 \times 56$ kbps $= 10.3$ kbps

$$\text{Transmission rate of stations} = \frac{1000 \text{ bits}}{100 \text{ sec}} = 10 \text{ bits/sec.}$$

- Let N be the number of stations that can use the channel.

$$\therefore \quad N = \frac{R}{10 \text{ bits/sec}}$$

$$= \frac{10.3 \text{ kbps}}{10} = 1030$$ ...Ans.

Ex. 4.7.6 : A slotted ALOHA network transmits 200-bit frames using a shared channel with a 200-kbps bandwidth. Find the throughput if the system (all stations together) produces 1) 1000 frames per second, 2) 500 frames per second, 3) 250 frames per second.

Soln. :

1. In slotted ALOHA for G = 1 maximum throughput = 0.368

Throughput $= 0.368 \times 1000 = 368$ frames/sec.

2. In this case G = 500 frames/sec. i.e. $\frac{1}{2}$

$$S = Ge^{-4} = \frac{1}{2}e^{-(1/2)} = 0.3032$$

Throughput $= 0.3032 \times 500$

$= 151.63 \cong 151$ frames/sec.

3. In this case G = 250 frames/sec. i.e. $\frac{1}{4}$

$$\therefore \quad S = Ge^{-4} = \frac{1}{4} \times e^{-(1/4)} = 0.1947$$

$\therefore$ Throughput $= 0.1947 \times 250$

$= 48.67 \cong 49$ frames/sec.

Review Questions

Q. 1 Explain the layered architecture of LAN explaining the function of the LLC and MAC sublayer.

Q. 2 What is static and dynamic channel allocation ?

s no need of CSMA/CD for a full
N ?
MA/CD.
MA/CA ?
MA.

Module 4

Chapter 5

Network Layer

Syllabus

Network Layer design issues, Communication Primitives: Unicast, Multicast, Broadcast.
Routing algorithms : Shortest Path (Dijkastra's), Link state routing, Distance Vector Routing.
Congestion control algorithms : Open loop congestion control, Closed loop congestion control, QoS parameters, Token and Leaky bucket algorithms.

Chapter Contents

5.1 Network Layer :

- The network layer is responsible for carrying the packet from the source all the way to destination.
- In short it is responsible for host-to-host delivery.
- The network layer has a higher responsibility than the data link layer, because the data link layer is only supposed to move the frames from one end of the wire to the other end.
- Thus network layer is the lowest layer that deals with the end to end transmission.

5.1.1 Position of Network Layer :

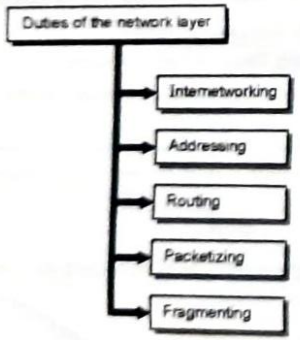
- Fig. 5.1.1 shows the position of network layer in the 5 layer internet model. It is the third layer.

(G-433) **Fig. 5.1.1 : Position of network layer**

- It receives services from the data link layer and provides services to the transport layer.

5.1.2 Network Layer Duties :

- Fig. 5.1.2 shows the set of duties of the network layer.



(G-434) **Fig. 5.1.2 : Duties of the network layer**

1. ***Internetworking :***

- This is the main duty of network layer. It provides the logical connection between different types of networks.

2. **Addressing :**

- Addressing is necessary to identify each device on the Internet uniquely.
- This is similar to a telephone system.
- The addresses used in the network layer should be able to uniquely define the connection of a computer to the Internet universally.

3. **Routing :**

- In a network, there are multiple roots available from a source to a destination and one of them is to be chosen.
- The network layer decides which root is to be taken.
- This is called as routing and it depends on various criterions.

4. **Packetizing :**

- As discussed earlier, the network layer receives the packets from upper layer protocol and encapsulates them to form new packets.
- This is called as packetizing. A network layer protocol called IP (Internetworking Protocol), does the job of packetizing.

5. **Fragmenting :**

- The sent datagram can travel through different networks.
- Each router decapsulates the IP datagram from the received frame.
- Then the datagram is processed and encapsulated in another frame.

Other issues :

- The other issues which are not directly related to the duties of network layer but need to be discussed are :
  1. Address resolution.
  2. Multicasting.
  3. Routing protocols.

Other supporting protocols :

- The Internetworking Protocol (IP) needs the support of another protocol ICMP or ARP etc. in the network layer.

How to achieve the goals ?

- In order to achieve the goals, the network layer must know about the topology of the communication subnet i.e. the set of all routers.
- It also should choose appropriate paths for communication.
- The routes should be chosen in such a way that overloading of some routers and idle operation of others should be avoided.

5.2 Network Layer Design Issues :

- The important network layer design issues include the service provided to the transport layer and the internal design of subnet.
- The network layer has been designed with the following goals :
  1. The services provided should be independent of the underlying technology.
  2. Users of the service need not know about the physical implementation of the network.
- This design goal has great importance because there is a great variety of networks in operation.
- The design of the layer must not disable us from connecting to networks of different technologies.
- The transport layer (that is the host computer) should be shielded from the number, type and different topologies of the subnets user uses.
- That is, all that transport layer wants is a communication link, it need not know how that link is established.
- Finally, there is a need for some uniform addressing scheme for network addresses.
- With these goals in mind, two different types of services emerged :
  1. Connection oriented Network Services
  2. Connectionless Network Services.
- **A connection-oriented service** is one in which the user is given a "reliable" end to end connection.
- To communicate, the user first makes a request for connection, then uses the connection to communicate his content, and then closes the connection.
- A telephone call is the classic example of a connection oriented service.
- In ***a connectionless service,*** the user simply puts his information into bundles called packets, puts an address on it, and then sends it for the destination.
- There is no guarantee that the bundle will reach the destination.
- So a connectionless service is one which is similar to the postal system.

5.2.1 Store and Forward Packet Switching :

- Refer Fig. 5.2.1 which demonstrates the environment of the network layer protocols.

(G-435) **Fig. 5.2.1 : The environment of the network layer protocols**

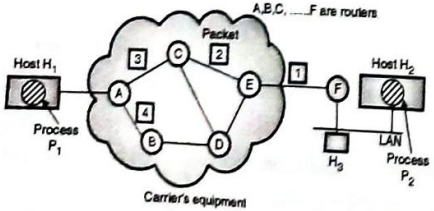
- This system of Fig. 5.2.1 is made up of following components :
  1. Carrier equipments (routers and transmission lines).
  2. Customer's equipments.
- H_1 is host – 1 and it is directly connected to router A via a leased line.
- Host H_2 is on a LAN which is connected to router F.
- Host H_1 wants to send a packet. So it communicates with its nearest router (A).
- Router A will store the packet until it has fully arrived so that the checksum can be verified.
- Then the packet is forwarded to the next router (B).
- This process continues till it reaches the destination host H_2.
- This mechanism is called as the store and forward packet switching.

5.2.2 Services Provided to the Transport Layer :

- The network layer services are designed to achieve the following goals :
  1. The services should not be dependent on the subnet technology.
  2. Transport layer should not be exposed to the number, type and topology of the subnet.
  3. The network address which is made available to the transport layer must use a uniform numbering plan.
- The network service can be connectionless or connection oriented.
- The Internet has a connectionless network layer whereas the ATM networks have a connection oriented network layer.
- The connection oriented and connectionless services both have their own sets of advantages and disadvantages.
- Finally we can say that the network layer should provide a raw means to send packets from a to b and that is all.

5.2.3 Implementation of Connectionless Service :

- In the connectionless service, the packets from sending host H_1 are injected into the subnet individually and each packet is routed independently as shown in Fig. 5.2.2.

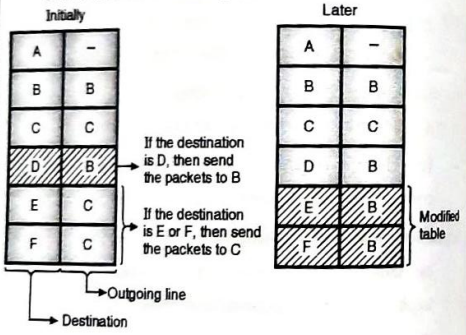


(G-436) Fig. 5.2.2 : Routing within a datagram subnet

- No advanced connection establishment is required.
- The packets are called as **datagrams** and the subnet is called as **datagram subnet.**

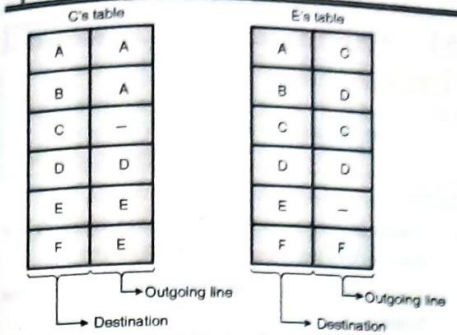
Working :

- Process P_1 on host H_1 wants to send a long message to process P_2 on host H_2.
- Let this message be broken into four packets 1, 2, 3 and 4 at the network layer.
- Then all these packets are sent to router A.
- Every router has its internal table which tells it where to send packets for each possible destination.
- Each entry in the router's table is a pair that consists of a destination and the outgoing line to be used to send the packet for that destination.
- In Fig. 5.2.2, C has two outgoing lines E and D.
- So every packet coming to router C should be sent to either D or E, even if the ultimate destination is F.
- Fig. 5.2.3(a) shows the routing table of A. It has two tables named as **initially** and **later.**



(G-437) Fig. 5.2.3(a) : Routing tables of A

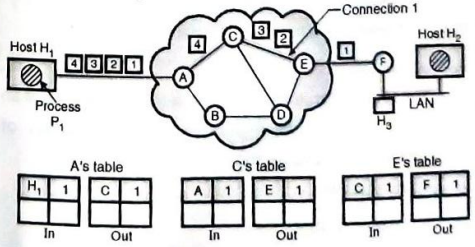
- As per the **initial** routing table of A, since the destination is F the packets 1, 2 and 3 were first sent to C, then to E and finally to F.
- But when packet 4 arrived at the input of A, even though the destination was F, the packet was not sent to C instead it was sent to B.
- The reason can be a traffic jam along the ACE path.
- As soon as A learned about the traffic jam along the ACE path it modified its routing table as shown in Fig. 5.2.3(a) as **later** and routed the 4th packet via path ABDEF.
- Fig. 5.2.3(b) shows the routing tables for C and E.



(G-438) Fig. 5.2.3(b) : Routing tables of C and E

5.2.4 Implementation of Connection-Oriented Service :

- For the connection oriented service, a path from source to destination needs to be established before sending any data packet.
- This connection is called as **Virtual Circuit (VC)** and the subnet is called as the **Virtual Circuit Subnet.**
- Here all the packets will follow the same path which was established before communication.
- When the connection is opened, the virtual circuit is also terminated.
- In the connection oriented service, each packet carries an identifier.
- This identifier can tell us about the virtual circuit (VC) that this packet belongs to.
- Refer Fig. 5.2.4. Host H_1 has established connection 1 with host H_2.



(G-439) Fig. 5.2.4 : Routing within a VC subnet

- This connection is remembered as the first entry in each routing table.
- As shown in Fig. 5.2.4, the first line of A's table shows that if a packet having connection identifier 1 arrives from H_1, it should be routed to C and a connection identifier 1 should be given to it.
- Similarly the first line of C's table shows that it routes the packets to E with an identifier 1.

5.2.5 Internal Organization of the Network Layer :

- Basically there are two philosophies for organizing the subnet :
  1. To use connection oriented service.
  2. To use connectionless service.
- In the connection oriented service, a connection is called as **virtual circuit.**
- It is similar to a physical connection between the sender host and the destination host.
- In the connectionless organization, the independent packets are called as **datagrams.** They are analogus to telegrams.

Virtual circuits :

- The principle behind the virtual circuits is to choose only one route from source to destination.
- When a connection is established, it is used for sending all the traffic over this connection.
- When the connection is released, the virtual circuit is terminated.

Datagram :

- With a datagram, the routes from source to destination are not decided in advance.
- Each packet sent is routed independently. Different packets of the same message can follow different routes.
- The datagram subnets have to do more work but they are more robust and deal with failures and congestion more easily as compared to virtual circuit subnets.

Features of virtual circuits :

- In virtual circuits every router will have to maintain and update a table.
- Each packet must have a virtual circuit number field in its header in addition to sequence number checksum etc.
- It is necessary to setup a VC before communication.

- The users are charged for connect time as well as for the amount of data transported.

Features of a datagram :

- The routers do not have to maintain any tables.
- Each datagram must contain full destination address. These addresses can be very long.
- When a packet comes in, the router finds an available outgoing line and sends the packet out on that line.
- So that it can reach the destination.

5.2.6 Comparison of Virtual Circuit and Datagram Subnets :

MU : Dec. 09, May 10, May 11, May 12

University Questions

Q. 1 Differentiate between virtual-circuit and datagram subnets. **(Dec. 09, May 10, May 11, 10 Marks)**

Q. 2 Compare virtual circuits and datagram subnets and show their diagrammatic representation during congestion control. **(May 12, 10 Marks)**

- Table 5.2.1 shows the comparison of VC subnet and datagram subnets.

Table 5.2.1 : Comparison of VC and Datagram Subnets

Sr. No.	Parameter	VC subnet	Datagram subnet
1.	Connection set up	Required	Not required.
2.	Addressing	Each packet contains a short VC number	Each packet contains the source as well as destination address.
3.	Repairs	Harder to repair	Easy to repair.
4.	State information	A table is needed to hold the state information.	Subnet does not hold state information.
5.	Routing	Route chosen is fixed. All packets follow this route. This is static routing.	Each packet is routed independently. This is dynamic routing.
6.	Congestion control	Easy	Difficult.
7.	Effect of router failure.	All VCs which passed through failed router are terminated.	No other effect except for the packets lost at the time of crash.

5.3 Routers :

MU : May 04

University Questions

Q. 1 What does routing mean and how does it works ? Describe routing table structure. **(May 04, 10 Marks)**

Definition :

- Router is a networking device which forwards the data packets between computer networks.
- Routers perform the traffic directing function on the Internet.

Functions :

- The two important functions, performed by a router are :
  1. Determination of path (routing)
  2. Packet forwarding.

OSI layer :

- Routers work at the network layer of the OSI model.
- Routers are devices that connect two or more networks.
- They consist of a combination of hardware and software.
- The hardware can be in the form of a network server, a separate computer or a special device, as well as the physical interfaces to the various networks in the internetwork.

5.3.1 Routing Tables :

MU : May 04

University Questions

Q. 1 What does routing mean and how does it works ? Describe routing table structure. **(May 04, 10 Marks)**

- The routing table for a host or a router consists of an entry for each destination, or a combination of destinations to route the IP packets.
- Routing tables can be of two types :
  1. Static routing tables
  2. Dynamic routing tables

1. Static routing table :

- The information in the static routing tables is entered manual. The route of a packet to each destination is entered into the table by the administrator.

- This routing table cannot update itself automatically.
- It has to be changed manually as and when required.
- Hence static routing table is useful only for small networks.

2. Dynamic routing table :

- The dynamic routing tables can get automatically updated by using a dynamic routing protocol such as RIP, OSPF or BGB.
- The structure of a dynamic routing table is shown in Table 5.3.1.

Table 5.3.1 : Format of dynamic routing table

Mask	Network Address	Next hop address	Interface	Flags	Reference count	Use

5.4 Routing :

MU : May 04

University Questions

Q. 1 What does routing mean and how does it works ? Describe routing table structure. **(May 04, 10 Marks)**

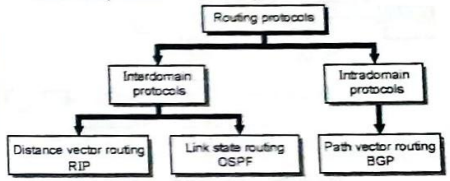
- Routing is a very important issue in the network layer.
- A router creates its routing table so as to help forwarding a datagram in the connectionless services.
- It also helps in creating a virtual circuit in the connection oriented service.
- In the following sections we are going to discuss about the types of routing and different routing algorithms such as distance vector routing, link state routing and hierarchical routing.

5.4.1 Types of Routing :

- Routing can be broadly classified into three types :
  1. Unicast routing.
  2. Broadcast routing
  3. Multicast routing.
- We can also classify the routing into two types as follows :
  1. Intra-domain routing.
  2. Inter-domain routing.

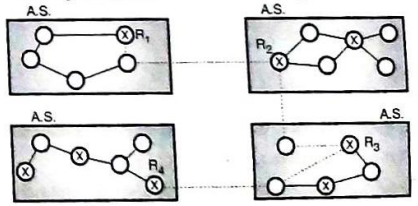
5.4.2 Intra and Inter-domain Routing :

- Today the size of the Internet is so big that one routing protocol cannot handle the task of updating the routing tables of all the routers.
- Hence an internet is divided into **Autonomous Systems** (AS).
- An Autonomous System (AS) is a group of networks and routers which is controlled by a single administrator. An AS is shown in Fig. 5.4.1.
- The **intra-domain routing** is defined as the routing inside an autonomous system whereas the routing between autonomous system is known as the **inter-domain routing.**
- Several intradomain and interdomain protocols are used. They are as shown in Fig. 5.4.1.



(G-1291) Fig. 5.4.1 : Classification of routing protocols

- The examples of interdomain routing protocols are :
  1. Distance vector routing
  2. Link state routing.
- An example of intradomain routing protocol is path vector routing.
- Each A.S. is allowed to choose one or more intradomain routing protocols in order to handle the routing inside the A.S. as shown in Fig. 5.4.2.



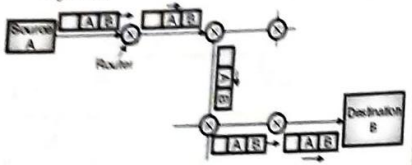
(G-1292) Fig. 5.4.2 : Autonomous systems

- But only one interdomain routing protocol will handle routing between autonomous systems.
- The Routing Information Protocol (RIP) is an implementation of distance vector routing.

- Whereas the OSPF is an implementation of link state protocol.
- The BGP is an implementation of the path vector protocol.

5.4.3 Unicast Routing :

- In unicast routing there is a one to one relation between the source and the destination.
- That means only one source sends packets to only one destination.
- The type of source and destination addresses included in the IP datagram are unicast addresses assigned to the hosts.
- The concept of unicast routing is illustrated in Fig. 5.4.3.



(G-448) **Fig. 5.4.3 : Unicast routing**

- In unicast routing when a router receives a packet, it forwards that packet through only one of its ports which corresponds to the optimum path.
- The router can discard the packet if it can not find the destination address.

Metric :

- A metric is defined as the cost assigned for passing through a network.
- The metric assigned to each network depends on the type of protocol.

Interior and exterior routing :

- An Internet is so large that for one routing protocol it is impossible to handle the task of updating the routing tables of all the routers.
- So an Internet is divided into a number of Autonomous Systems (AS). An AS is group of networks and routers.

Interior routing :

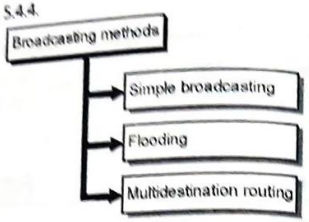
- The routing that takes place inside an AS is called as interior routing.

Exterior routing :

- The routing that takes place among various autonomous systems is called as exterior routing.

5.4.4 Broadcast Routing :

- In certain applications, the host has to send packets to many or all other hosts.
- If the sender send a packet to all destinations simultaneously then it is called as **broadcasting.**
- Various methods of broadcasting are as shown in Fig. 5.4.4.



(G-449) **Fig. 5.4.4 : Various methods of broadcasting**

1. **Simple broadcasting :**
   - In this method the source will simply send a distinct (a separate) packet to each destination.
   - This method has two drawbacks :
     1. A lot of bandwidth is wasted.
     2. The source has to have a complete list of all destinations.
2. **Flooding :**
   - Flooding is another method used for broadcasting.
   - The problem with flooding is that it has a point to point routing algorithm.
   - So it consumes a lot of bandwidth and generates too many packets.
3. **Multi-destination routing :**
   - This is the third algorithm used for broadcasting.
   - In this algorithm each packet will contain a list of destinations or a bit map which indicates the desired destination.
   - When such a packet arrives at a router, the router first checks all the destinations.
   - Then it decides the set of output lines that will be required based on the destination addresses.
   - The router then generates a new copy of the received packet for each output line to be used.

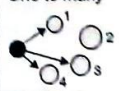
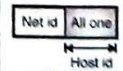
- It includes a list of only those destinations that are to use the line in each packet going out on that line.
- This will save bandwidth to a great extent. Also generation of too many packets right from the sending end will also be avoided.

5.4.5 Multicast Routing :

- In multicasting a message from a sender is to be sent to a group of destinations but not all the destinations in a network.
- A process has to send a message to all other processes in the group.
- For a small group it is possible to send a point-to-point message.
- But this is expensive if the group is large. So we have to send messages to a well defined groups which are small compared to the network size.
- Sending message to such a group is called **multicasting** and the routing algorithm used for multicasting is **multicast routing.**
- Multicast routing is a special class of broadcast routing.

5.4.6 Comparison of Broadcast and Multicast Routing :

Table 5.4.1 : Comparison of broadcast and multicast

Sr. No.	Basis of comparison	Broadcast	Multicast
1.	Addressing	A broadcast addressing the type of address that identifies all nodes except the sending one.	A multicast addressing uses one to many type association between sender and receiver.
2.	Concept	One to all.  (G-2514)	One to many  (G-2514)
3.	IPv4 address format	Broadcast address  (G-2515)	Multicast address  (G-2515)
4	IPv6	Does not support broadcasting	Supports multicasting.

5.5 Routing Algorithms :

- One of the important functions of the network layer is to route the packets from the source machine to the destination machine.
- The major area of network layer design includes the algorithms which choose the routes and the data structures which are used.
- **Routing algorithm** is a part of network layer software.
- It is responsible for deciding the output line over which a packet is to be sent.
- Such a decision is dependent on whether the subnet is a virtual circuit or it is datagram switching.

5.5.1 Desired Properties of a Routing Algorithm :

- There are certain desirable properties of a routing algorithm as follows :
  1. Correctness
  2. Robustness
  3. Stability
  4. Fairness and
  5. Optimality.

5.5.2 Types of Routing Algorithms :

MU : May 07, Dec. 08, May 09, Dec. 12, May 13, Dec. 15

University Questions

Q. 1 What are the different types of Routing Algorithms ? Explain any one in detail. **(May 07, May 13, 10 Marks)**

Q. 2 Explain different types of routing algorithm. **(Dec. 08, May 09, 10 Marks)**

Q. 3 What are the different types of routing ? Explain distance vector routing. **(Dec. 12, 10 Marks)**

Q. 4 What are the different types of routing algorithms ? When would we prefer to use hierarchical routing over link state routing ? **(Dec. 15, 10 Marks)**

- Routing algorithms can be divided into two groups :
  1. Non-adaptive algorithms.
  2. Adaptive algorithms.

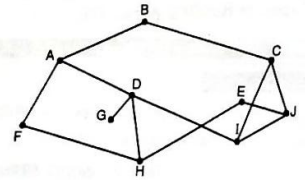
1. **Non-adaptive algorithms :**
   - For this type of algorithms, the routing decision is not based on the measurement or estimation of current traffic and topology.
   - However the choice of the route is done in advance, off-line and it is downloaded to the routers.
   - This is called as static routing.
2. **Adaptive algorithms :**
   - For these algorithms the routing decision can be changed if there are any changes in topology or traffic etc.
   - This is called as dynamic routing.
   - In the following sections we are going to discuss various static and dynamic algorithms.

5.5.3 Optimality Principle :

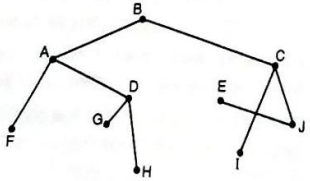
- A general statement about optimality is called as optimality principle.
- It states that if router J is on the optimal path from router I to router K, then the optimal path from J to K will also be along the same route.

Sink tree :

- A set of optimal routes from all the sources to a given destination form a tree called sink tree and it is shown in Fig. 5.5.1. The root of the sink tree is at the destination.



(a) A subnet



(b) A sink tree for router B

(G-450) Fig. 5.5.1

- Note that a sink tree need not be unique.
- Other trees with the same path lengths may also exist.
- All the routing algorithms are supposed to discover and use the sink trees for all routers.
- In the sink tree of Fig. 5.5.1, the distance metric is the number of hops. In Fig. 5.5.1(b) a sink tree for router B has been shown.
- The paths from B to every router with minimum number of hops.

5.6 Static Algorithms :

MU : Dec. 04, May 05, Dec. 05, May 13

University Questions

Q. 1 What are the different types of routing algorithms ? Explain any one in detail. **(Dec. 04, May 05, May 13, 10 Marks)**

Q. 2 What is static routing ? What are advantages of dynamic routing ? Explain shortest path routing in detail ? **(Dec. 05, 5 Marks)**

- The examples of static algorithms are :
  1. Shortest path routing.
  2. Flooding.
  3. Flow based routing.

5.6.1 Shortest Path Routing :

MU : Dec. 05

University Questions

Q. 1 What is static routing ? What are advantages of dynamic routing ? Explain shortest path routing in detail ? **(Dec. 05, 5 Marks)**

- This algorithm is based on the simplest and most widely used principle.
- Here a graph of subnet is prepared in which each node represents either a host or a router and each arc represents a communication link.
- So as to choose a path between any two routers, this algorithm simply finds the shortest path between them.

How to decide the shortest path ?

- One way of measuring the path length is the number of hops.
- Another way (metric) is the geographical distance in kilometres.
- Some other metrics are also possible.

- For example we can label each arc (link) with the mean queuing and transmission delay and obtain the shortest path as the fastest path.

Labels on the arcs :

- The labels on the arcs can be computed as a function of distance bandwidth, average traffic, mean queue length, cost of communication, measured delay etc.
- The algorithm compares various parameters and calculates the shortest path, on the basis of any one or combination of criterions stated above.

Various shortest path algorithms :

- There are many algorithms for computing the shortest path between two nodes.
- One of them is Dijkstra algorithm. The other one is Bellman-Ford algorithm.

5.6.2 Dijkstra's Algorithm :

MU : May 10, Dec. 11

University Questions

Q. 1 Explain Dijkstra's algorithm as shortest path routing with example. **(May 10, Dec. 11, 10 Marks)**

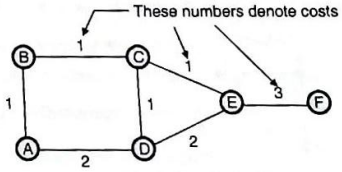
- Dijkstra's algorithm is used for computing the shortest path from the **root** node to every other node in the network.
- The root node is defined as the node corresponding to the router where the algorithm is being run.
- The total number of nodes are divided into two groups namely the P group and T group.
- In the P group we have those nodes for which the shortest path has already been found.
- In T group the remaining nodes are placed. The path to every node in the T group should be computed from a node which is already present in group P.
- We should find out every possible way to reach an outside node by a one hop path from a node which is already present in P and choose the shortest of these paths as the path to the desired node.
- As stated earlier we define two sets P (permanent) and T (temporary) of the nodes.
- In set P we have nodes to which the shortest path has already been found and in set T we have nodes to which we are considering the shortest paths.
- At the time of starting, P is initialized to the current node and T is initialized to null.

Steps for Dijkhatra's algorithm :

- The algorithm then repeats the following steps :

1. Start from the desired node say p. Write p in the P set.
2. For this node p, add each of its neighbours n to T set. The addition of these nodes in T will have to satisfy the following conditions :
   1. If the neighbouring node (say n) is not there in T then add it annotating it with the cost to reach it through p and p's ID.
   2. If n is already present in T and the path to n through p has a lower cost, then remove the earlier instance of n and add the new instance annotated with the cost to reach it through p and p's ID.
   3. Pick up the neighbour n which has the smallest cost in T, and if it is not present in P then add it to P. Use its annotation to determine the router p to use to react n.
   4. Stop when T is empty.
- This algorithm will be clear after solving the following example.

Ex. 5.6.1 : For the network shown in Fig. P. 5.6.1(a), show the computations at node A using the Dijkshtra's algorithm.



(G-451) Fig. P. 5.6.1(a) : Given network

Soln. :

Step 1 :

- Since the computations are to be done at node A, the starting node will be A.
- We enter this node into group P as shown in Table P. 5.6.1(a).
- We add the neighbouring nodes B and D in group T alongwith the costs to reach them through A as shown in Table P. 5.6.1(a).

(G-451(a)) Table P. 5.6.1(a)

Permanent (P)	Temporary (T)
A	B(A,1),D(A,2)

Note : B(A,1) means B is reached by A and the cost is 1. Similarly D(A,2) means D is reached by A and the cost is 2.

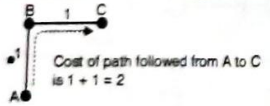
Step 2 :

- Now pick up the neighbour with the smallest cost and add it to P set. Here the neighbour with smallest cost is B.
- So let us add B(A,1) to P group as shown in Table P. 5.6.1(b).
- As B is added to P group, we have to add its neighbour i.e. C to the T group, as shown in Table P. 5.6.1(b).

(G-452) Table P. 5.6.1(b)

Permanent (P)	Temporary (T)
A	B(A,1),D(A,2)
A,B(A,1)	D(A,2),C(B,2)

- Note that D(A,2) has remained in T group as it is but C(B,2) is a new entry. C(B,2) means C is reached by A via B with a cost of 2.
- The cost is 2 due to the path followed from A to B and then to C, as illustrated in Fig. P. 5.6.1(b).



(G-453) Fig. P. 5.6.1(b)

Step 3 :

- Now pick up the neighbour in T set with the smallest cost in Table P. 5.6.1(b) and add it to the P set.
- Here we choose neighbour D because it is the immediate neighbour of A.
- Since D is added to P group, we have to add its neighbours i.e. C and E to the T group as shown in Table P. 5.6.1(c).

(G-454) Table P. 5.6.1(c)

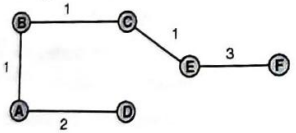
Permanent (P)	Temporary (T)
A	B(A,1),D(A,2)
A,B(A,1)	D(A,2),C(B,2)
A,B(A,1),D(A,2)	E(D,4),C(B,2)

- Note that C(B,2) goes as it is, and E(D,4) is a new entry to Table P. 5.6.1(c).
- But C(D,3) can not be entered because its cost is 3.
- Where E(D,4) means E is reached by A via D and the cost is 4.
- Similarly we can proceed further. The final table is as shown in Table P. 5.6.1(d).

(G-455) Table P. 5.6.1(d) : Final table

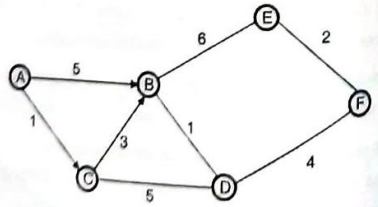
Permanent (P)	Temporary (T)
A	B(A,1),D(A,2)
A,B(A,1)	D(A,2),C(B,2)
A,B(A,1) D(A,2)	E(D,4),C(B,2)
A,B(A,1) D(A,2),C(B,2)	E(C,3) E(D,4) can not be included
A,B(A,1) D(A,2),C(B,2) E(C,3)	F(E,6) F(E,7) can not be included
A,B(A,1) D(A,2),C(B,2) E(C,3),F(E,6)	Empty (NULL)

- The shortest paths from A to all other nodes are as shown in Fig. P. 5.6.1(c).



(G-456) Fig. P. 5.6.1(c) : Shortest paths from A to all other nodes

Ex. 5.6.2 : For the network shown in Fig. P. 5.6.2(a) show the computations at node A using the Dijkshtra's algorithm.



(G-457) Fig. P. 5.6.2(a) : Given network

Soln. :

Step 1 :

- The starting node is A. Enter it in to group P as shown in Table P. 5.6.2(a).
- Add the neighbours B and C to the temporary group T.

(G-457(a)) Table P. 5.6.2(a)

Permanent (P)	Temporary (T)
A	B(A,5),C(A,1)

Step 2 :

- Now pick up the neighbour with smallest cost i.e. C and add it to group P.
- As C is added to P group, we have to add D i.e. the neighbour of C to the T group as shown in Table P. 5.6.2(b).

(G-458) Table P. 5.6.2(b)

Permanent (P)	Temporary (T)
A	B(A,5),C(A,1)
A, C(A,1)	B(A,5),D(C,6),B(C,4)

- B(C, 4) is another entry in T group which shows that B is approached by A via C and the cost is 4.

Step 3 :

- Now move B(C,4) from T to P group and add neighbours E and D to the T group as shown in Table P. 5.6.2(c).

(G-459) Table P. 5.6.2(c)

Permanent (P)	Temporary (T)
A	B(A,5),C(A,1)
A, C(A,1)	D(C,6),B(C,4)
A, C(A,1),B(C,4)	D(C,6),E(B,10)

- Note that E(B,10) corresponds to the route A-C-B-E with a cost 1 + 3 + 6 = 10.
- Do not use the route A-B-E because the associated cost is 5 + 6 = 11.

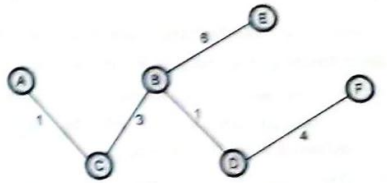
Step 4 :

- Now continue in the same manner to get the final table as shown in Table P. 5.6.2(d).

(G-460) Table P. 5.6.2(d) : Final table

Permanent (P)	Temporary (T)
A	B(A,5),C(A,1)
A, C(A,1)	D(C,6),B(C,4)
A, C(A,1),B (C,4)	D(C,6),E(B,10)
A, C(A,1) B(C,4), D(C,6)	E(B, 10) F(D, 10)
A, C(A,1) B(C,4), D(C,6) E(B,10)	F (D, 10)
A, C(A,1) B(C,4), D(C,6) E(B,10), F(D,10)	Null (Stop)

- The shortest path from A to other nodes is shown in Fig. P. 5.6.2(b).



(G-461) Fig. P. 5.6.2(b) : Shortest paths from A to all other nodes

- Dijkstra's algorithm is most suitable for the dense networks and it is particularly useful for the parallel implementation, i.e. when the scan operation is carried out in parallel.
- The disadvantages are that it does not take any advantage of sparsity well and it is only appropriate for the networks with positive arc lengths.

5.6.3 Flooding :

- This is another static algorithm.
- In this algorithm every incoming packet is sent out on every outgoing line except the line on which it has arrived.
- That is why the name flooding. Each line except the incoming lines are flooded with the copies of the same packet.
- One disadvantage of flooding is that it generate a large number of duplicate packets.
- In fact it produces infinite number of duplicate packets unless we somehow stop the process.

- There are various damping techniques such as :
  1. Using a hop counter.
  2. To keep a track of which packets have been flooded.
  3. Selective flooding.
- To prevent endless copies of packets circulating for very long time through the network a hop count may be used to suppress onwards transmission of packets after a number of hops which exceed the network "diameter".
- The other problem is that destination must be prepared to receive multiple copies of an incoming packet.
- Flooding has two interesting characteristics that arise from the fact that all possible routes are tried :
  1. As long as there is a route from source to destination the packet will be definitely delivered to the destination.
  2. One copy of the packet will reach the destination via the quickest possible route.

Selective flooding :

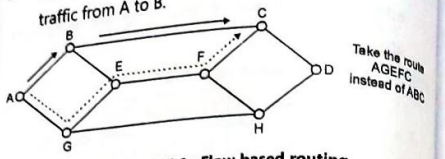
- This is slightly more practical type of flooding principle.
- In this algorithm every incoming packet is not sent out on every output line.
- Instead packet is sent only on those lines which are likely to go in the desired direction.

Applications of flooding :

- Flooding does not have many practical applications.
- But it is useful in military applications where a large number of routers are blown into pieces (damaged) at any instant.
- So placing a packet on every outgoing line really makes sense.
- In such applications robustness of flooding is very much desirable.
- Second application is in the distributed database applications.
- Flooding always chooses the shortest path so it produces the shortest possible delay.

5.6.4 Flow Based Routing :

- This is a static algorithm which uses topology and load condition (traffic) for deciding a route.
- For example in Fig. 5.6.1, there is always a huge traffic from A to B.



(G-462) **Fig. 5.6.1 : Flow based routing**

- Then the traffic from A to C should not be routed through B.
- Instead route it through AGEFC even though it is a longer path than ABC. This is called as a **flow based routing**.
- It is possible to optimise the routing by analysing the data flow mathematically.
- This is possible if the average traffic from one node to the other is known in advance and it is constant in time.
- The mathematical analysis is based on idea that for a given line if the capacity and average data flow are known, then it is possible to calculate the mean packet delay using the Queueing theory.
- From the mean delays on all the lines it is possible to calculate the mean packet delay for the whole subnet.
- To use the technique of flow based routing, the following information should be known in advance :
  1. Subnet topology
  2. Traffic matrix.
  3. Line capacity matrix which specifies capacity of each line.

5.7 Dynamic Routing Algorithms :

MU : Dec. 05

University Questions

Q. 1 What is static routing ? What are advantages of dynamic routing ? Explain shortest path routing in detail ? **(Dec. 05, 5 Marks)**

- The modern computer networks normally use the dynamic routing algorithms.

- Two dynamic routing algorithms namely distance vector routing and link state routing are used popularly.
- Both these algorithms are suitable for the packet switched networks.
- Both these algorithms assume that a router knows the address of each neighbouring router and the cost of reaching each neighbour.
- In the distance vector routing, each node tells its neighbours about its distance to every other node in the network.
- In the link state routing, a node tells every other node in the network the distance to its neighbours.
- So both these routing algorithms are distributed type and so they are suitable for large internetworks.

5.8 Distance Vector Routing Algorithm :

MU : May 11, May 12, Dec. 12, May 17, Dec. 18, May 19, Dec. 19

University Questions

Q. 1 Explain distance vector routing and its count to infinity problem. **(May 11, 10 Marks)**

Q. 2 Explain DVR algorithm and mention the drawbacks of the algorithm when put into practice. **(May 12, 10 Marks)**

Q. 3 What are the different types of routing ? Explain distance vector routing. **(Dec. 12, 10 Marks)**

Q. 4 Explain distance vector routing. What are its limitations and how are they overcome ? **(May 17, 10 Marks)**

Q. 5 Write a short note on: Distance Vector Routing **(Dec. 18, 5 Marks)**

Q. 6 Explain distance vector routing protocol. What is count to identify problem How to overcome it ? **(May 19, 10 Marks)**

Q. 7 Explain the difference between static and dynamic routing. Explain distance vector routing in detail. **(Dec. 19, 10 Marks)**

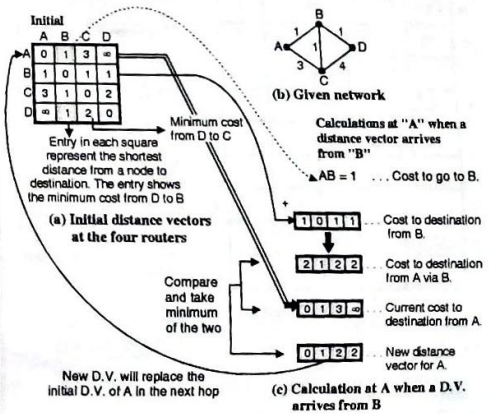
- In this algorithm, each router maintains a table called vector, such a table gives the best known distance to each destination and the information about which line to be used to reach there.
- This algorithm is sometimes called by other names such as :
  1. Distributed Bellman-Ford routing algorithm.
  2. Ford-Fulkerson algorithm
- In distance vector routing, each router maintains a routing table.
- It contains one entry for each router in the subnet.
- This entry has two parts :
  1. The first part shows the preferred outgoing line to be used to reach the specific destination.
  2. Second part gives an estimate of the time or distance to that destination.

Distance vector :

- In distance vector routing, we assume that each router knows the identity of every other router in the network, but the shortest part to each router is not known.
- A **distance vector** is defined as the list of <destination, cost> tuples, one tuple per destination. Each router maintains a distance vector.
- The cost in each tuple is equal the sum of costs on the shortest path to the destination.

Updation of router tables :

- A router periodically sends a copy of its distance vector to all its neighbours.
- When a router receives a distance vector from its neighbour, it tries to find out whether its cost to reach any destination would decrease if it routed packets to that destination through that particular neighbouring router. This is illustrated in Fig. 5.8.1.

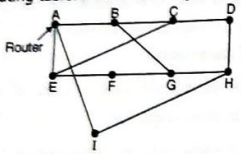


(G-463) **Fig. 5.8.1 : Distance vector algorithm at router A**

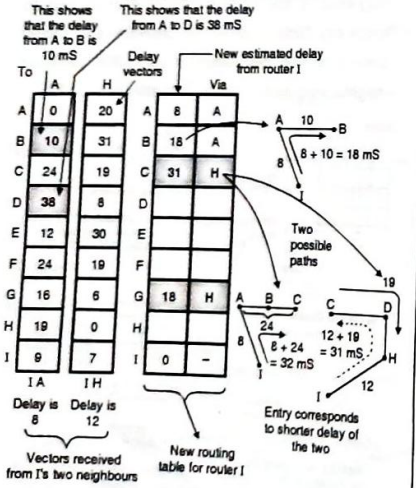
- Fig. 5.8.1 shows how the D.V. at A is automatically modified when a D.V. is received from B.
- A similar calculation takes place at the other routers as well. So the entries at every router can change.
- In Fig. 5.8.1(a) the initial distance vector is shown.
- The entries indicate to the costs corresponding to the shortest distance between the routers indicate to that square.
- For example, AC = 3 indicates the cost corresponding to the shortest path in terms of number of hops from A to C.
- Even if nodes asynchronously update their distance vectors the routing tables eventually converge.
- The well known example of distance vector routing is the Bellman-Ford algorithm.

Routing procedure in distance vector routing :

- The example of a subnet is shown in Fig. 5.8.2(a) and the routing tables are shown in Fig. 5.8.2(b).

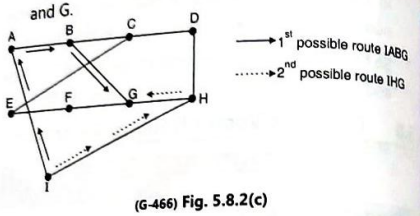


(G-464) **Fig. 5.8.2(a) : A subnet**



(G-465) **Fig. 5.8.2(b) : Routing table**

- The entries in router tables of Fig. 5.8.2(b) are the delay vectors.
- For example consider the shaded boxes of Fig. 5.8.2(b).
- The entry in the first shaded box shows that the delay from A to B is 10 msec, whereas the entry in the other shaded box indicates that the delay from A to D is 38 msec.
- Consider how router I computes its new route to router G.
- Fig. 5.8.2(c) shows the two possible routes between I and G.



(G-466) **Fig. 5.8.2(c)**

- I knows that the reach G via A, the delay required is : (L-891)

I to A Delay = 8mS
A to G Delay = 16mS
$\therefore$ I to G Delay = 8 + 16 = 24 msec

- Whereas the delay between I and G via H (route IHG) is : (L-892)

I to H Delay = 12mS
H to G Delay = 6mS
$\therefore$ I to G Delay = 12 + 6 = 18 msec

- The best of these values is 18 msec corresponding to the path IHG.
- Hence it makes an entry in its routing table (I's table) that the delay to G is 18 msec and that the route to use it is via H.
- The new routing table for router I is shown in Fig. 5.8.2(b).
- Similarly we can calculate the delays, from I to different destinations from A to I and enter the minimum possible delay into the I's router table.

5.8.1 Disadvantages :

1. The distance vector routing takes a long time in converging to the correct answer.
   - This is due to a problem called count-to-infinity problem.

- This problem can be solved by using the split horizon algorithm.

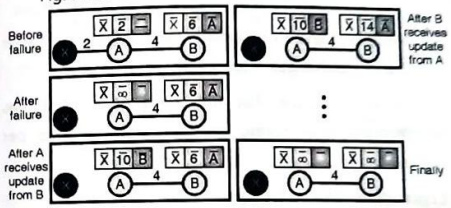
2. Another problem is that this algorithm does not take the line bandwidth into consideration when choosing a root.
   - This is a serious problem due to which this algorithm was replaced by the Link State Routing algorithm.

5.8.2 Looping in Distance Vector Routing Protocol :

- A problem in distance vector rooting is its instability. A network using this protocol can become unstable.

Two node loop instability :

- A network with three nodes has been shown in Fig. 5.8.3.



(G-1499) **Fig. 5.8.3 : Two node loop instability**

- Note that the routing tables are shown partially for discussion.
- At the beginning both nodes A and B know how to reach node X.
- But the link joining A and X fails suddenly. So node A changes its table.
- If A could send its changed routing table to B immediately, everything is okay. No problem will occur.
- But the system becomes unstable if B sends its routing table to A before receiving A's routing table.
- This is because node A receives the updated B's routing table and assumes that B has found a new path to reach node X.
- So A immediately updates its routing table (which is incorrect).
- Based on this update now A sends its new update to B.
- Now B thinks that something has changed around A and so it updates its routing table.
- Due to this process, the cost of reaching X increases gradually and finally becomes infinite.
- At this moment both A and B understand that now it is impossible to reach X.
- Note that during this entire time the system is unstable.
- A thinks that the route to X goes via B whereas B thinks that the route is via node A.
- So if A receives a packet for X, it goes to B and then again returns back to A.
- Similarly if B receives a packet destined for X, it goes to A and returns back to B.
- This bouncing of packets between nodes A and B is known as the **two-node loop problem.**
- This problem can be solved by using one of the following strategies :
  1. Defining infinity
  2. Split horizon
  3. Split horizon and poison reverse.
- There is a similar problem called three node loop problem present in the system using distance vector routing.

5.8.3 Count to Infinity Problem :

MU : May 11, May 15, May 17, May 19

University Questions

Q. 1 Explain distance vector routing and its count to infinity problem. **(May 11, 10 Marks)**

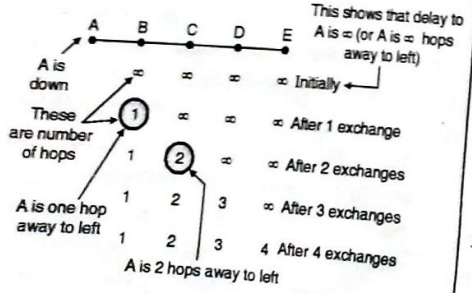
Q. 2 What is count to infinity problem in distance vector routing ? Discuss in detail. **(May 15, 10 Marks)**

Q. 3 Explain distance vector routing. What are its limitations and how are they overcome ? **(May 17, 10 Marks)**

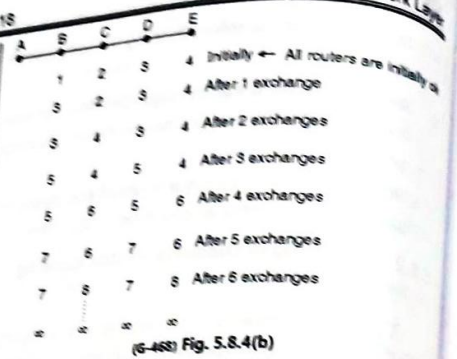
Q. 4 Explain distance vector routing protocol. What is count to identify problem How to overcome it ? **(May 19, 10 Marks)**

- Theoretically the distance vector routing works properly but practically it has a serious problem.
- The problem is that we get a correct answer but we get it slowly.
- In other words it reacts quickly to good news but it reacts too slowly to bad news.

- Consider a router whose best route to destination X is large.
- If on the next exchange neighbour A suddenly reports a short delay to X, the router will switch over and start using the line to A for sending the traffic to destination X.
- Thus in one vector exchange, the good news is processed.
- Let us see how fast does a good news propagate. Consider a linear subnet of Fig. 5.8.4 which has five nodes. The delay metric used is the number of hops.
- Assume that A is initially down and that all the other routers know this.
- So all the routers have recorded that the delay to A is infinity.
- When A becomes OK, the other routers come to know about it via the vector exchanges.
- Then suddenly a vector exchange at all the routers will take place simultaneously.
- At the time of first vector exchange, B comes to know that its left neighbour has a zero delay to A.
- So as shown in Fig. 5.8.4(a), B makes an entry in its routing table that A is one hop away to the left.
- All the other routers still think that A is down. So in the second row of Fig. 5.8.4(a), the entries below C D E are ∞.
- On the second vector exchange, C comes to know that B has a path of 1 hop length to A, so C updates its routing table and indicates a path of 2 hop length.
- But D and E do not change their table entries.

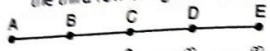


(G-467) Fig. 5.8.4(a)



(G-468) Fig. 5.8.4(b)

- So after the second vector exchange the entries in the third row of Fig. 5.8.4(a) are :

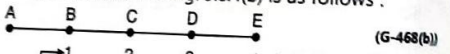


(G-468(a))

- Similarly D and E will update their routing tables after 3 and 4 exchanges respectively.
- So we conclude that the good news of A has recovered has spread at a rate of one hop per exchange.

Explanation of Fig. 5.8.4(b) :

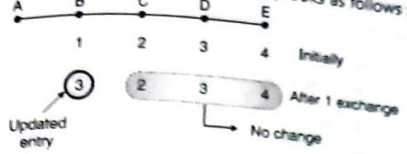
- Now refer Fig. 5.8.4(b). Here initially all routers are OK.
- The routers B, C, D and E have distances of 1, 2, 3 and 4 respectively to A.
- So the first row of Fig. 5.8.4(b) is as follows :



(G-468(b))

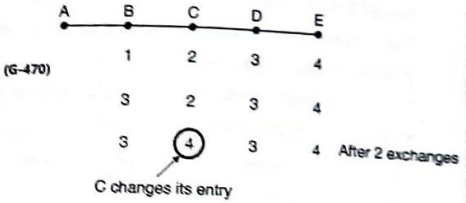
- Now imagine that suddenly A goes down or line between A and B is cut.
- At the first packet exchange B does not hear anything from A (because A is down).
- But C says " I have a path of length 2 to A". But poor B does not understand that this path is through B itself.
- So B thinks that it can reach A via C with a path length 3. (B to C 1 hop and C to A 2 hops) so it accordingly updates its routing table. But D and E do not update their entries.

- So the second row of Fig. 5.8.4(b) looks as follows :



(G-469)

- On the second exchange C realizes that both its neighbours (B and D) claim to have a path of length 3 to A.
- So it picks one of them at random and makes its new distance to A as 4.
- This is shown in row 3 of Fig. 5.8.4(b). It is repeated below.



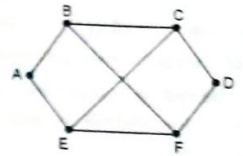
(G-470)

- Similarly the other routers keep updating their tables after every exchange.
- It is expected that finally we should get ∞ in the router tables of B, C, D and E indicating that A is down.
- We do reach this state at the end in Fig. 5.8.4(b) but after a very long time.
- The conclusion is bad news propagates slowly. This problem is called as **count-to-infinity** problem.
- The solution to this problem is to use the split horizon algorithm.

5.8.4 Split Horizon Algorithm :

- To avoid the count to infinity problem, several changes in the algorithm have been suggested.
- But none of them work satisfactorily in all situations.
- One particular method which is widely implemented, is called as the **split horizon algorithm.**
- In this algorithm, the minimum cost to a given destination is not sent to a neighbour if the neighbour is the next node along the shortest path.
- For example if node A thinks that the best route to node B is via node C, then node A should not send the corresponding minimum cost to node C.

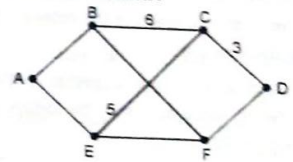
Ex. 5.8.1 : Consider the network of Fig. P. 5.8.1. Distance vector routing is used, and the following vectors have just come in to router C : From B : (5, 0, 8, 12, 6, 2); from D : (16, 12, 6, 0, 9, 10); and from E : (7, 6, 3, 9, 0, 4). The cost of the links from C to B, D and E, are 6, 3 and 5 respectively. What is C's new routing table ? Give both the outgoing line to use and the cost.



(L-924) Fig. P. 5.8.1

Soln. :

Step 1 : Redraw the network :



(L-919) Fig. P. 5.8.1(a)

Step 2 : Vectors received by C from B, D, E are as follows :

(L-920) Table P. 5.8.1

To	B	D	E
A	5	16	7
B	0	12	6
C	8	6	3
D	12	0	9
E	6	9	0
F	2	10	4

- Therefore C's routing table is as follows.

(L-921) Table P. 5.8.1(a) : C's routing table

To	Cost	Next
A	11	B
B	6	-
C	0	-
D	3	-
E	5	-
F	8	B

5.9 Link State Routing :

MU : Dec. 13, May 16, Dec. 16

University Questions

Q. 1 What are the steps involved in link state routing. Explain the contents and the requirements of link state packets. **(Dec. 13, May 16, Dec. 16, 10 Marks)**

- Distance vector routing was used in ARPANET upto 1979.
- After that it was replaced by the link state routing.
- Variants of this algorithm are now widely used.
- The link state routing is simple and each router has to perform the following five operations.

Router operations :

1. Each router should discover its neighbours and obtain their network addresses.
2. Then it should measure the delay or cost to each of these neighbours.
3. It should construct a packet containing the network addresses and the delays of all the neighbours.
4. Send this packet to all other routers.
5. Compute the shortest path to every other router.

- The complete topology and all the delays are experimentally measured and this information is conveyed to each and every router.
- Then a shortest path algorithm such as Dijkshtra's algorithm can be used to find the shortest path to every other router.

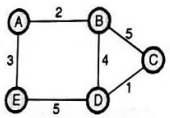
Protocols :

- *Link state routing is popularly used in practice.*
- *The OSPF protocol which is used in the Internet uses the link state algorithm.*
- *IS-IS i.e.* Intermediate system – Intermediate system *is the other protocol which uses the link state algorithm.*
- *IS-IS is used in Internet backbones and in some digital cellular systems such as CDPD.*

Building a routing table in link state routing :

Link state routing :

- Now we will discuss the development of routing table in link state routing.
- Here the term **link state** is used for defining the characteristic of a link or edge, which represents a network in the Internet.
- The **cost** associated with each link is important. The links having lower costs are preferred to the links having higher costs.
- A nonexisting or broken link is indicated by an ∞ cost.
- In this method, each node must have a complete map of the network.
- That means each node should have complete information about the state of each link.
- The collection of states of all the links in an Internet is called as **Link-State Database (LSDB).**
- For the entire Internet, there is only one LSDB and its copy is available with each node.
- Each node uses it to create the least cost tree. The example of LSDB is as shown in Fig. 5.9.1(b) for the Internet shown in Fig. 5.9.1(a).



(a) Internetwork

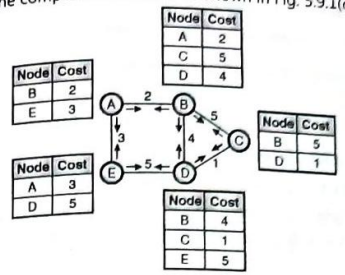
	A	B	C	D	E
A	0	2	∞	∞	3
B	2	0	5	4	∞
C	∞	5	0	1	∞
D	∞	4	1	0	5
E	3	∞	∞	5	0

(b) Link state database (LSDB)

(G-2201) Fig. 5.9.1

- The next step is creation of LSDB (which contains all the information about the Internet) at each node.
- This can be achieved by a process called **flooding.** Each node sends a greeting message to all its immediate neighbours, so as to collect two important pieces of information as follows :
  1. The identity of the neighbouring node.
  2. Cost of the link.
- The packet containing this information is called as **LS Packet** (LSP), which is sent out of each interface.

- After receiving all the new LSPs each node will create the comprehensive LSDB as shown in Fig. 5.9.1(c).



(G-2202) Fig. 5.9.1(c)

- This LSDB is same for each node which shows the whole map of the internet.
- That means a node can use the LSDB to make the whole map of the Internet.

5.10 Comparison of Link State Routing and Distance Vector Routing :

Table 5.10.1 : Comparison of Link State Routing and Distance Vector Routing

Sr. No.	Distance vector routing	Link state routing
1.	Each router maintains routing table indexed by and containing one entry for each router in the subnet.	It is the advanced version of distance vector routing
2.	Algorithm took too long to converge.	Algorithm is faster.
3.	Bandwidth is less.	Wide bandwidth is available.
4.	Router measure delay directly with special ECHO packets.	All delays measured and distributed to every router.
5.	It doesn't take line bandwidth into account when choosing the routes.	It considers the line bandwidth into account when choosing the routes.

5.11 Hierarchical Routing :

MU : Dec. 13, Dec. 15

University Questions

Q. 1 What are the advantages and disadvantages of hierarchical routing ? **(Dec. 13, 5 Marks)**

Q. 2 What are the different types of routing algorithms ? When would we prefer to use hierarchical routing over link state routing ? **(Dec. 15, 10 Marks)**

- As the size of the network increases, the size of the routing tables of the routers also increases.
- As a result of large routing tables, the router memory is consumed to a great extent, more CPU time is needed to scan the tables and more bandwidth is required to send status report about the tables.
- Sometimes the network becomes so large that the size of the router table becomes excessively large and practically it becomes impossible for every router to have an entry for all the other routers except itself.
- Then the hierarchical routing such as the one used in telephone networks should be used.
- In this type of routing the total number of routers are divided into different **regions.**
- A router will know everything about the all other belonging to its own region only.
- It does not know anything about the internal structure of other regions. This reduces the size of the router table.
- When various networks are connected together, each network is treated as a separate region.
- For very large networks the hierarchy is prepared as follows :

  Level 1 : Regions

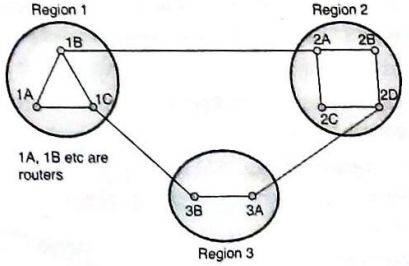
  Level 2 : Clusters : it is a group of regions.

  Level 3 : Zones : zone is a group of clusters.

  Level 4 : Groups : group contains many zones.

Two Level Hierarchical Routing :

- For networks of smaller size, a two level hierarchical routing is sufficient.
- Fig. 5.11.1(a) shows network containing 3 regions. Fig. 5.11.1(b) shows the full routing table of router 1A which has 9 entries because in all there are 9 routers.



(G-471) Fig. 5.11.1(a) : A network

Full routing table for 1A

Destination	Line	Hops
1 A	–	–
1 B	1 B	1
1 C	1 C	1
2 A	1 B	2
2 B	1 B	3
2 C	1 B	3
2 D	1 B	4
3 A	1 C	3
3 B	1 C	2

(G-2304) Fig. 5.11.1(b) : Full routing table for router 1A

- Now with a two level hierarchical routing, the routing table of the same router reduces to a much smaller size as shown in Fig. 5.11.1(c). This table has only 5 entries.

Hierarchical routing table for 1A

	Destination	Line	Hops
Region 1	1 A	–	–
	1 B	1 B	1
	1 C	1 C	1
Region 2 → 2		1 B	2
Region 3 → 3		1 C	2

(G-2305) Fig. 5.11.1(c) : Hierarchical routing table for router 1A

- In the hierarchical table of Fig. 5.11.1(c), there are entries for all local routers (1 A, 1 B and 1 C) belonging to the region of 1 A as before. But there are no detailed entries for the other regions.
- Instead all other regions have been compressed into a single router per region.
- For example traffic from 1A to any router in region-2 is via 1 B-2 A line as shown by the shaded entry in Fig. 5.11.1(c).
- Similarly all the traffic from 1A to region 3 is routed through the line 1C-3B.
- Comparison of Figs. 5.11.1(b) and (c) shows how hierarchical routing reduces the size of routing tables.

Disadvantage :

- The reduced table size has a price tag attached to it.
- It comes at the expense of increased path length.
- But it is practically acceptable.

How many levels a hierarchy should have ?

- Kamoun and Kleinrock have discovered that for an N router subnet, the optimum number of hierarchy levels is $\log_e N$ and it requires a total of $\log_e N$ entries per router table.

Ex. 5.11.1 : For hierarchical routing with 4800 routers, what region and cluster sizes should be chosen to minimize the size of the routing table for a three-layer hierarchy ?

Soln. :

- The three level hierarchy has got the three levels as shown in the following diagram. **(L-910)**

Level 3 : Clusters
|
Level 2 : Regions
|
Level 1 : Routers

- If the number of clusters is x, number of regions per cluster is y, and the number of routers in each region is z then the each router needs z entries for the local routers, (y – 1) entries for routing to other regions within its own cluster and (x – 1) entries for distant clusters.

∴ Total number of entries in the router table

$= (x-1) + (y-1) + z = x + y + z - 2$

5.11.1 Comparison of Static and Dynamic Routing : MU : Dec. 19

University Questions

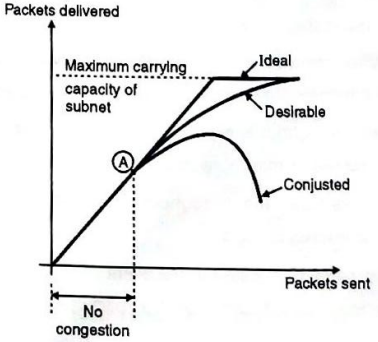
Q. 1 Explain the difference between static and dynamic routing. Explain distance vector routing in detail. **(Dec. 19, 10 Marks)**

Table 5.11.1 : Comparison of static and dynamic routing

Sr. No.	Parameter	Static routing	Dynamic routing
1.	Updating of the routing tables	Manually done	Automatically done
2.	Bandwidth requirement	Less	More
3.	Application area	In small networks	In large networks
4.	Routing protocols	None	EIGRP, ARP etc.
5.	Security	Highly secure	Less secure
6.	Routing algorithms	Shortest path, flooding, flow based routing	Distance vector, link state
7.	Link failure	Any link failure affects the other routing paths	Does not affect other routing paths
8.	Additional resources	Not required	Required to store information
9.	Routing decision	Not based on the measured or estimated current traffic	Is based on the changes in topology or traffic

5.12 Congestion Control :

- An important issue in a packet switching network is congestion.
- If an extremely large number of packets are present in a part of a subnet, the performance degrades. This situation is called as congestion.
- Congestion in a network may occur when the load on the network i.e. the number of packets sent to the network is greater than the capacity of the network (i.e. the number of packets a network can handle).
- Fig. 5.12.1 explains the concept of congestion graphically.



(G-473) Fig. 5.12.1 : Concept of congestion

- Upto point A in Fig. 5.12.1, the number of packets sent into the subnet by the host is within the capacity of the network.
- So all these packets are delivered. In short the number of packets delivered is proportional to number of packets sent and no congestion takes place.
- But after point A, the traffic increases too far. The routers cannot cope with the increased traffic and they begin to lose packets. The congestion begins here.
- As the traffic increases further, the performance degrades more and more packets are lost and congestion worsens.
- At very high traffic, the performance collapses completely and almost all packets are lost. This is the worst possible congestion.

5.12.1 Need of Congestion Control : MU : Dec. 14, May 19

University Questions

Q. 1 Why there is a need for congestion control ? What are the different mechanisms ? Explain them. **(Dec. 14, 10 Marks)**

Q. 2 Explain congestion control. Explain leaky bucket algorithm. **(May 19, 10 Marks)**

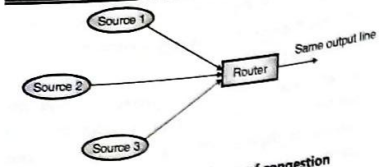
- It is not possible to completely avoid the congestion but it is necessary to avoid it otherwise control it.
- Congestion will result in long queues, which results in buffer overflow and loss of packets.
- So congestion control is necessary to ensure that the user gets the negotiated QoS (Quality of Service).

5.12.2 Causes of Congestion :

- Some of the causes of congestion are as follows :

1. Sudden flow of packets from multiple senders :

- 1If suddenly a flow of packets starts coming on three or four senders all needing the same output line.
- Then a queue will become long. If the memory capacity is not sufficient to hold all these packets, some of them will be lost.
- This is shown in Fig. 5.12.2(a).

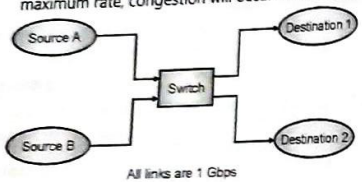


(G-474) **Fig. 5.12.2(a) : Cause of congestion**

- This leads to congestion. Note that increasing the memory to infinity also does not solve the problem, in fact it worsens.

2. Slow and low bandwidth links :

- The problem will be solved when high speed links become available.
- It is not always the case, sometimes increases in link bandwidth can aggravate the congestion problem because higher speed links may make the network more unbalanced.
- For the configuration shown in Fig. 5.12.2(b), if both the sources begin to send to destination 1 at their maximum rate, congestion will occur at the switch.



(G-475) **Fig. 5.12.2(b) : Network with high speed links**

- Higher speed links can make the congestion condition in the switch worse.

3. Slow processors :

- Congestion is caused by slow processors. The problem will be solved when processor speed is improved.
- Faster processors will transmit more data in unit time.
- If several nodes begin to transmit to one destination simultaneously at their maximum rate, the destination will be overwhelmed soon.

4. Multiple transmission of same packets :

- Congestion can make itself worse. If a router does not have any free buffers it should ignore (discard) new packets arriving at it.
- But when a packet is discarded, the sender may retransmit it many times because it is not receiving the acknowledgement of the packet.
- This multiple transmission of packets will force the congestion to take place at the sending end.

5.12.3 Effects of Congestion :

- Due to congestion, various network parameters get affected.
- The most important network parameters which get affected by congestion are as follows :
  1. Delay
  2. Throughput

1. Delay versus load :

- When the load is very small as compared to the capacity of network, the delay is at its minimum.
- This minimum delay is equal to the sum of propagation delay and processing delay and it is negligible.
- But when load increases and becomes comparable to the network capacity the delay increases sharply due to the addition **waiting time** which gets added to the total delay.
- The delay becomes infinite when the load is greater than the capacity of network.
- The delay increases due to congestion.

2. Throughput versus load :

- Throughput can be defined as number of packets passing through a network per unit time.
- When the load is less than the capacity of network, the throughput increases proportional to the load.
- The throughput is expected to remain constant after it reaches its maximum (capacity of network).
- But practically the throughput reduces sharply as the load reaches its capacity.
- This happens because in the event of congestion, the routers discard some of the packets received by them.
- Thus throughput reduces due to congestion.

5.12.4 Difference between Congestion Control and Flow Control : MU : Dec. 08

University Questions

Q. 1 Compare congestion control and flow control. **(Dec. 08, 5 Marks)**

- Congestion control makes it sure that the subnet is able to carry the offered traffic i.e. the subnet is able to carry all the packets sent by all the senders to their destinations.
- Congestion control is dependent on the behaviour of all the hosts, all the routers and other factors which reduce the carrying capacity of a subnet.
- On the contrary, the flow control is related to point to point traffic between a sender and its destination.
- Flow control ensures that a fast sender does not send data at a rate faster than the rate at which the receiver can receive it.
- Flow control involves some kind of feedback from the receiver, which can control the sending rate of the sender.

5.12.5 Principle of Congestion Control :

MU : Dec. 05, Dec. 15, May 16, Dec. 16, Dec. 17, Dec. 18

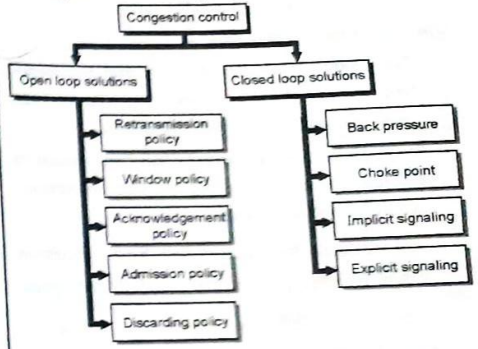
University Questions

Q. 1 Write short notes on : Congestion control. **(Dec. 05, 8 Marks)**

Q. 2 What is congestion control ? Explain various congestion prevention policies. **(Dec. 15, Dec. 17, 10 Marks)**

Q. 3 Compare open loop congestion control, closed loop congestion control. **(May 16, Dec. 16, Dec. 18, 10 Marks)**

- The solutions to the congestion problems can be divided into two categories or groups as open loop solutions and closed loop solutions.
- Congestion control refers to the techniques and mechanisms which can either prevent congestion from happening or remove congestion after it has taken place.
- The **open loop congestion** control is based on the prevention of congestion whereas the **closed loop solutions** are for removing the congestion after it has occurred.
- Fig. 5.12.3 shows the classification of congestion control schemes and various policies used in open loop and closed loop groups.



(G-476) **Fig. 5.12.3 : Classification of congestion control schemes**

Open loop control :

- Open loop solutions try to solve the congestion issue by excellent design to prevent the congestion from happening.
- Open loop control is exercised by using the tools such as deciding when to accept the new packets, when to discard the packets, which packets are to be discarded and making the scheduling decisions at various points.
- It is important to note that none of these decisions are made on the basis of the current status of a network, as no feedback is being used.

Closed loop control :

- The closed loop congestion control uses some kind of feedback. It takes into account the current status of the network.
- A closed loop control is based on the following three steps :
  1. Detect the congestion and locate it by monitoring the system.
  2. Transfer the information about congestion to places where action can be taken.
  3. Adjust the system operations to correct the congestion.

- Two examples of closed loop control are :
  1. TCP flow control.
  2. BR rate control for an ATM network.

Open loop versus closed loop :

- Open loop approaches do not need end-to-end feedback, one of the examples of this type are prior-reservation and hop-to-hop flow control.
- In closed-loop approaches, the source can adjust its cell rate on the basis of the feedback information received from the network.
- Some people feel that closed loop congestion control schemes are too slow in todays high-speed, large range network.
- Because it takes a long time for feedback to go back to source.
- Hence before any corrective action takes place thousands of packets have been already lost.
- But on other hand, if the congestion has already taken place and the overload is of long duration, the congestion cannot be released unless the source causing the congestion is asked to reduce its rate.
- Furthermore, ABR service is designed to use any bandwidth that is left over the source must have some knowledge of what is available when it is sending cells.

5.12.6 Congestion Prevention Policies :

MU : May 12, Dec. 12, Dec. 15, May 17, Dec. 17

University Questions

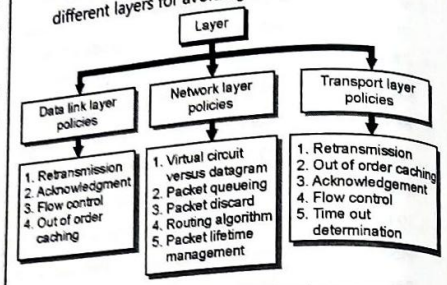
Q. 1 List the design features to be taken care of as congestion prevention policies in the different layers of network. **(May 12, 10 Marks)**

Q. 2 What are the congestion prevention policies ? Explain the congestion control in virtual circuit and datagram subnets. **(Dec. 12, 10 Marks)**

Q. 3 What is congestion control ? Explain various congestion prevention policies. **(Dec. 15, Dec. 17, 10 Marks)**

Q. 4 What are congestion prevention policies ? Explain congestion control in virtual circuit and datagram subnets. **(May 17, 10 Marks)**

- In this section we are going to discuss the open loop congestion control systems.
- These systems try to avoid congestion by using the appropriate policies at different levels.
- Fig. 5.12.4 lists various policies corresponding to different layers for avoiding congestion.



(G-477) Fig. 5.12.4 : Policies affecting the congestion

Policies related to data link layer :

1. Retransmission policy :

- The retransmission policy and the retransmission timers must be designed to optimise efficiency and at the same time prevent congestion.
- The retransmission policy deals with how fast a sender times out.
- If a sender times out early then it will retransmit all the packets and such a retransmission can lead to congestion.
- By designing the retransmission policy we can avoid this and prevent congestion.

2. Out of order caching policy :

- If the receivers routinely discard all the packets which are out of order, then retransmission of these packets will take place.
- This will increase the load and result in congestion.
- So a selective repeat (retransmission) should be adopted to avoid congestion.

3. Acknowledgement policy :

- If each received packet is promptly acknowledged then the acknowledgement packets will increase the traffic.

- If the acknowledgement is delayed (piggybacking) then there is a possibility of time out and retransmission.
- So a tight flow control has to be exercised to avoid congestion.

4. Window policy :

- The type of window at the sender may also affect congestion.
- The selective repeat window is better than the Go Back N window.

Policies related to network layer :

1. Choice between virtual circuit and datagrams :

- This choice at the network layer will affect the congestion because many congestion control algorithms work only with virtual circuit subnets.

2. Packet queuing and service :

- This policy is related to whether the routers have one queue per input line and one queue per output line or both.
- This policy is also related to the order in which the packets are processed e.g. round robin or priority based etc.

3. Discard policy :

- This policy lays a rule which tells the routers about which packet is to be discarded.
- A good discard policy can prevent congestion and a bad one will worsen the situation.

4. Routing algorithms :

- The routing algorithms can spread the traffic over all the lines.
- By doing so it is ensured that none of the lines are overloaded.
- This will certainly avoid congestion.

5. Package lifetime management :

- This policy decides the maximum time for which a packet may live before being discarded.
- This time should be of adequate value so that congestion can be avoided.

Policies related to transport layer :

- The policies at the transport layer are same as those at the data link layer.
- But at transport layer determining the time out interval is more difficult.
- If it is too short then extra packets are sent unnecessarily whereas if it is too long, congestion will reduce at the cost of increased response time (network will become slow).

Traffic shaping :

- One of the important reason behind congestion is the bursty nature of the traffic.
- If the traffic has a uniform data rate then congestion would not happen every now and then.
- But due to bursty traffic it can happen regularly.
- Traffic shaping is an open loop control. It prevents the congestion by making the packet transmission rate to be more predictable (bursty traffic is unpredictable).
- Thus traffic shaping will regulate the average rate or the burstiness of data transmission.
- Monitoring a traffic flow is called as **traffic policing**.
- Check if a packet stream (connection) is as per its descriptor, and if it is not as per its descriptor, then give penalty !
- In order to achieve this the network may want to monitor the traffic flow during the connection period.
- The process of monitoring and enforcing the traffic flow is called traffic policing.
- The types of penalties enforced are as follows :
  1. Drop packets that violate the descriptor.
  2. Give low priority to the packets violating the descriptor.

5.12.7 Congestion Control in Virtual Circuit Subnets :

MU : May 12, Dec. 12, May 17

University Questions

Q. 1 Compare virtual circuits and datagram subnets and show their diagrammatic representation during congestion control. **(May 12, 10 Marks)**

Q. 2 What are the congestion prevention policies ? Explain the congestion control in virtual circuit and datagram subnets. **(Dec. 12, 10 Marks)**

Q. 3 What are congestion prevention policies ? Explain congestion control in virtual circuit and datagram subnets. **(May 17, 10 Marks)**

- All the congestion control techniques discussed till now were open loop techniques.
- Now let us discuss a dynamic technique called **admission control.**

Admission control principle :

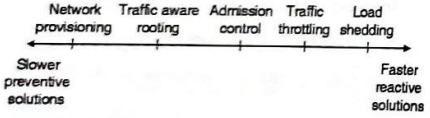
- This technique is used to keep the congestion which has already begun to a managable level and does not allow it to worsen any further.
- Its principle is as follows : Once congestion has been detected, do not set up any more virtual circuits until the congestion is cleared.
- The advantage of this technique is that it is a simple and easy to carry out control.

Alternative approach :

- An alternative approach to admission control allows the virtual circuits to set up even when a congestion has taken place.
- But carefully route all the new virtual circuits around the area where congestion is already present.

5.12.8 Approaches to Congestion Control :

- The two basic solutions to the problem of congestion are :
  1. Increase the resources 2. Decrease the load
- These solutions are applied on different time scales in order to either prevent congestion or handle it if it has occurred.



(G-1522) Fig. 5.12.5 : Time scales of approaches to congestion control

1. **Network provisioning :**

- The fundamental way of avoiding congestion is to build a network that is properly matched to the traffic that it is going to carry.
- If the network uses a low bandwidth link along which a heavy traffic is directed, then congestion is most certain to take place.
- We can add resources dynamically when there is congestion.
- For example, we can turn on additional routers and use spare (back up) lines whenever congestion has taken place.
- Another example is purchasing bandwidth on open market as and when congestion occurs.
- But you can't do it instantly. It takes a long time.
- This is called as **Network Provisioning.**
- It is a slow preventive solution and happens on a time scale of months.

2. **Traffic aware routing :**

- If we can not increase the capacity of a network then we should think of utilizing the existing capacity in the best possible why.
- Routers can be tailored to suite traffic patterns that change during the day as network users wake and sleep in different time zones.
- The traffic can be routed over those paths which have less traffic at that time.
- This is known as traffic aware rooting.

3. **Admission control :**

- Sometimes it is not possible to increase capacity.
- Then the only possible way to fight congestion is to **decrease the load.**
- As stated earlier, in the virtual circuit networks, new connections are not allowed once congestion has been detected.
- This is a feedback (closed loop) control approach.
- When the congestion is predicted, the network can deliver feedback to those sources who are responsible for congestion.
- Then these sources would be requested to **reduce** their outputs.
- There are two difficulties faced in this approach :
  1. It is difficult detect the of beginning of congestion.
  2. It is also difficult to inform the sources to slow down accordingly.
- The **leaky bucket** and **token bucket** methods are examples of admission control.

4. **Traffic throttling (Congestion avoidance) :**

- In the Internet and many other computer networks, senders adjust their transmission rates and send only that much traffic which a network can readily deliver without causing congestion.
- This is done so as to operates the network just before the beginning point of congestion.
- When congestion is about to happen the senders should be told to **reduce** their transmission and slow down.
- This technique is an example of **congestion avoidance** principle.
- The first step in traffic throttling is to **detect** the beginning point of congestion and the second step is to tell the senders to slow down.
- Note that traffic throttling approach can be used in both datagram **subnets** as well as **virtual circuit subnets.**
- The onset of congestion can be detected if the routers are made tomonitor the following parameters :
  1. Utilization of output links.
  2. Buffering of queued packets inside the router.
  3. Number of packets lost due to inadequate buffering.
- Generally the second parameter is most useful in practice.
- The second task for the routers is that they should deliver timely feedback to the senders.
- Different schemes use different feedback mechanisms. Some of them are as follows :
  1. Choke packets.
  2. Explicit Congestion Notification (ECN).
  3. Hop by Hop back pressure.

5. **Load shedding :**

- When all other solutions fail to contain congestion, the network has no option but to discard packets that can not be delivered.
- A good policy for selecting which packets to discard can help preventing the congestion collapse.

5.13 Congestion Control in Datagram Subnets :

MU : May 12, Dec. 12, Dec. 13, Dec. 14, May 17

University Questions

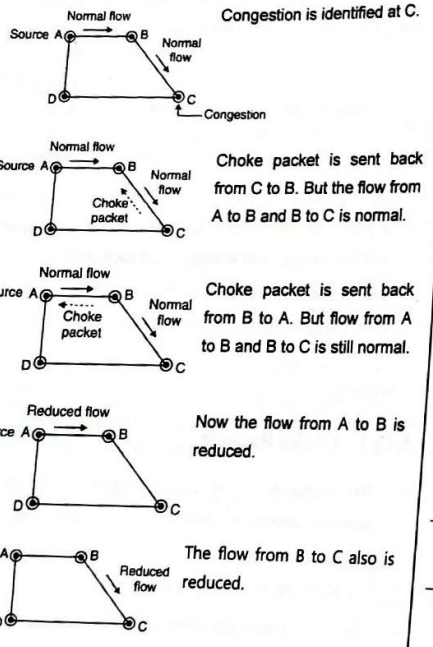
- **Q. 1** Compare virtual circuits and datagram subnets and show their diagrammatic representation during congestion control. **(May 12, 10 Marks)**
- **Q. 2** What are the congestion prevention policies ? Explain the congestion control in virtual circuit and datagram subnets. **(Dec. 12, 10 Marks)**
- **Q. 3** Explain the various methods for congestion control used in datagram subnets. **(Dec. 13, 10 Marks)**
- **Q. 4** Why there is a need for congestion control ? What are the different mechanisms ? Explain them. **(Dec. 14, 10 Marks)**
- **Q. 5** What are congestion prevention policies ? Explain congestion control in virtual circuit and datagram subnets. **(May 17, 10 Marks)**

- Let us now discuss some congestion control approaches which can be used in the datagram subnets (and also in virtual circuit subnets).
- The techniques are :
  1. Choke packets
  2. Load shedding
  3. Jitter control.

5.13.1 Choke Packets :

- This approach can be used in virtual circuits as well as in the datagram subnets.
- In this technique each router associates a real variable with each of its output lines.
- This real variable say "u" has a value between 0 and 1 and it indicates the how much is utilization of that line in percentage (60 %, 70 % etc.).
- If the value of "u" goes above the threshold then that output line will enter into a "warning" state.
- The router will check each newly arriving packet to see if its output line is in the "warning state".
- If it is in the warning state then the router will send back a **choke packet** signal to the sending host.
- The sender host will not generate any more data packets. This will reduce the congestion.

- Different congestion control algorithm have been proposed, depending on the value of thresholds.
- Depending on the threshold value, the choke packets can contain a mild warning, a stern warning or an ultimatum.
- Another algorithm may use the queue lengths or buffer utilization instead of using the line utilization as a deciding factor.
- The general concept of choke packet mechanism is demonstrated in Fig. 5.13.1(a).

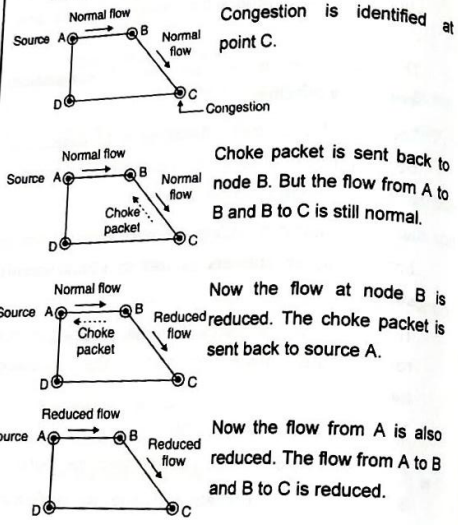


(G-478) Fig. 5.13.1(a) : Choke packet mechanism

- Fig. 5.13.1(a) shows that, the choke packets have to travel over the entire network, from the point of congestion to the appropriate source (i.e. from C to A).
- Then the action of reducing the flow will take place.
- The whole process is therefore very much time consuming.

Hop-by-Hop choke packet technique :

- The problem associated with the general choke packet mechanism can be overcome by using another technique called as hop-by-hop choke packet technique.
- This is demonstrated in Fig. 5.13.1(b). In this approach, the choke packets are used at each hop between the destination and source.



(G-479) Fig. 5.13.1(b) : Concept of hop-by-hop choke packet mechanism

- Each node receiving the choke packet will reduce its output flow.
- This will have a more effective and fast control over the overall transmission rate.
- Fig. 5.13.1(b) shows how the transmission rate is reduced at every hop in response to the choke packets.

Disadvantage :

- The problem with choke packet technique is that the action to be taken by the source host on receiving a choke packet is not compulsory.
- The host may reduce its transmission rate or ignore the choke packets.

Weighted fair queuing :

- The disadvantage of choke packet technique can be overcome with the help of the weighted queuing technique.
- The queuing algorithm was proposed first in 1987.
- In this algorithm it is proposed that the routers have a number of queues for each output line, with one queue for each source.

5.13.2 Load Shedding :

- Admission control, choke packets, fair queuing are the techniques suitable for light congestion.
- But if these techniques cannot eliminate the congestion, then the load shedding technique is to be used.
- The principle of load shedding states that when the routers are flooded with the packets that they cannot handle, they should simply throw the packets away.
- A router which is flooding with packets due to congestion can discard any packet at random. But there are better ways of doing this.
- The policy for dropping a packet depends on the type of packet.
- For file transfer an old packet is more important than a new packet.
- In contrast for multimedia a new packet is more important than an old one.
- Accordingly a policy is formulated for discarding the packets.
- An intelligent discard policy can be decided depending on the applications.
- It is not possible to implement such an intelligent discard policy without the, co-operation from the sender.
- The applications should mark their packets as per priority to indicate how important they are.
- If this is done then when the packets are to be discarded the routers can first drop packets having lower priority (i.e. the packets which are least important).
- Then the routers will discard the packets from next lower class and so on.
- One or more header bits are required to put the priority of a packet.
- In every ATM cell, 1 bit is reserved in the header for marking the priority.
- Every ATM cell is labeled either as a low priority or high priority.

5.13.3 Jitter Control : MU : Dec. 13

University Questions

Q. 1 Explain the various methods for congestion control used in datagram subnets. **(Dec. 13, 10 Marks)**

Definition of jitter:

- The delay introduced by the data communication networks is not constant.
- It varies packet to packet. The jitter measures the variability in packet delays and it is measured in terms of the difference of the minimum delay and maximum value of delay.
- Jitter is defined as the variation in delay for the packets belonging to the same flow.
- The real time audio and video cannot tolerate jitter on the other hand the jitter does not matter if the packets are carrying any data information contained in a file.
- For the audio and video transmission if the packets take 20 msec to 30 msec (delay) to reach the destination, it does not matter, provided that the delay remains constant.
- The quality of sound or video will be hampered if the delays associated with different packets have different values.

Jitter control :

- When a packet arrives at a router, the router will check to see whether the packet is behind or ahead and by what time.
- This information is stored in the packet and updated at every hop.
- If the packet is ahead of the schedule (early) then the router will hold it for a slightly longer time and if the packet is behind the schedule (late), then the router will try to send it out as quickly as possible.
- This will help in keeping the average delay per packet constant and will avoid time jitter.

5.13.4 Difference between End to End Delay and Jitter :

End to end delay :

- End to end delay is the time required for the signal to travel from transmitter to receiver.
- This delay is due to the time required for buffering, queueing, switch and routing.
- This time delay remains the same for all the types of packets in the same flow.

Jitter :

- Jitter is defined as the variation in delay for the packets belonging to the same flow.
- This is the difference between end to end delay and Jitter.

5.14 Quality of Service (QoS) :

MU : May 10, Dec. 10, May 11, May 13, Dec. 13

University Questions

Q. 1 Write short notes on : Quality of Service (QoS) of internetworking. **(May 10, May 11, 5 Marks)**

Q. 2 Discuss the quality of service requirements for audio on demand. **(Dec. 10, 5 Marks)**

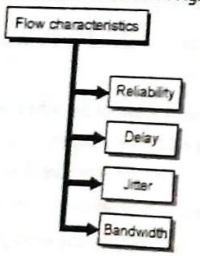
Q. 3 Explain the different factors associated with quality of service in internetwork. **(May 13, 10 Marks)**

Q. 4 Write short notes on : QoS requirements. **(Dec. 13, 10 Marks)**

- The long form of QoS is quality of service and it is an internetworking issue.
- We can informally define quality of service as something flow seeks to attain.

Flow characteristics/ QoS parameters :

- There are four important characteristics of data flow : reliability, delay, jitter and bandwidth.
- These characteristics are shown in Fig. 5.14.1.



(G-480) **Fig. 5.14.1 : Flow characteristics / QoS parameters**

1. **Reliability :**
   - A data flow must have some level of reliability. Lack of reliability means a packet or acknowledgment, will be lost and retransmission will be required.
   - However, each application programs has a different demand for reliability.
   - For example, it is more important that electronic mail, file transfer, and Internet access have reliable transmissions than telephony or audio conferencing.
2. **Delay :**
   - Source-to-destination delay is another important flow characteristic.
   - Again delay tolerance of different applications will be different.
   - In this case, telephony, audio conferencing, video conferencing, and remote log-in need minimum delay, while file transfer or email are delay tolerant applications.
3. **Jitter :**
   - Jitter is the variation in delay for packets belonging to the same flow. i.e. different packets experience different amounts of delays.
   - Real-time audio and video cannot tolerate a large amount of jitter.
   - On the other hand, it does not matter if packet carrying information in a file have different delays.
   - The transport layer at the destination waits until all packets arrive before delivery to the application layer.
4. **Bandwidth :**
   - Different applications need different bandwidths.
   - In video conferencing needs a huge bandwidth whereas an email may not need a large bandwidth.

5.14.1 Techniques for Achieving Good QoS :

- Some of the techniques useful in achieving good QoS are as follows :
  1. Buffering
  2. Traffic shaping
  3. Leaky bucket algorithm
  4. Token bucket algorithm
  5. Resource reservation
  6. Admission control
  7. Proportional routing
  8. Packet scheduling.

5.14.2 Traffic Shaping :

MU : Dec. 15, Dec. 18, Dec. 19

University Questions

Q. 1 What is traffic shaping ? Explain leaky bucket algorithm and compare it with token bucket algorithm. **(Dec. 15, Dec. 18, Dec. 19, 10 Marks)**

- One of the important reason behind congestion is the bursty nature of the traffic.
- If the traffic has a uniform data rate then congestion problem will not be very common.
- Traffic shaping is an open loop control of congestion control.
- It manages the congestion by making the packet transmission rate to be more predictable.
- This will make the data rate more uniform and bursty traffic is reduced.
- Thus traffic shaping will regulate the average rate or the burstiness of data transmission.
- The process of monitoring a traffic flow is called as **traffic policing.**
- Here the principle followed is to check if a packet stream (connection) obeys the rules and if it violates the rules then, give penalty !
- For this the network would like to monitor the traffic flow during the connection period.
- The process of monitoring and enforcing the rules to regulate traffic flow is called **traffic policing.**
- Penalty for breaking the rules will be :
  1. Drop packets that violate the rules.
  2. Give low priority to them.

Definition of traffic shaping :

- Traffic shaping is defined as a mechanism to control the amount and rate of the traffic sent to the network.

Traffic shaping techniques :

- The two popularly used traffic shaping techniques are :
  1. Leaky bucket algorithm and
  2. Token bucket algorithm

5.14.3 Leaky Bucket Algorithm :

MU : May 10, Dec. 15, Dec. 18, May 19, Dec. 19

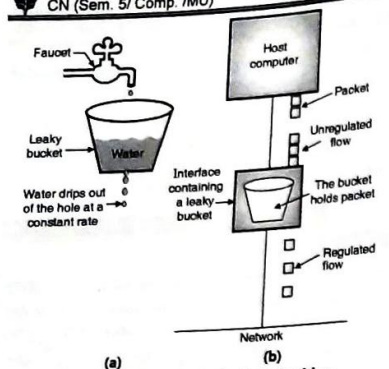
University Questions

Q. 1 Explain the working of leaky bucket algorithm. Why leaky bucket algorithm is used in Computer Network. **(May 10, 10 Marks)**

Q. 2 What is traffic shaping ? Explain leaky bucket algorithm and compare it with token bucket algorithm. **(Dec. 15, Dec. 18, Dec. 19, 10 Marks)**

Q. 3 Explain congestion control. Explain leaky bucket algorithm. **(May 19, 10 Marks)**

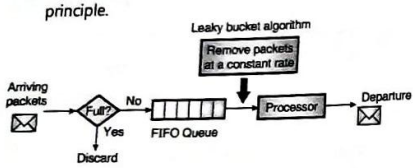
- Leaky bucket algorithm is used to control congestion in network traffic.
- As the name suggests it's working is similar to a leaky bucket in real life.
- The principle of leaky bucket algorithm is as follows :
- Leaky bucket is a bucket with a hole at bottom.
- Flow of the water from bucket is at a constant rate (data rate is constant) which is independent of water entering the bucket (incoming data).
- If bucket is full, any additional water entering in the bucket is thrown out (Packets are discarded).
- Same technique is applied to control congestion in network traffic.
- Every host in the network is having a buffer (equivalent to a bucket) with finite queue length.
- Packets which are put in the buffer when buffer is full are thrown away.
- The buffer may send some number of packets per unit time onto the subnet (helpful if packets vary greatly in size) as shown in Fig. 5.14.2.
- the data flow at the input of the bucket is unregulated but that at the bucket output is a regulated one.



(G-481) **Fig. 5.14.2 : Leaky bucket algorithm**

Leaky bucket implementation :

- Fig. 5.14.3 shows the implementation of leaky bucket principle.



(G-482) **Fig. 5.14.3 : Implementation of leaky bucket**

- A FIFO (First In First Out) queue is used for holding the packets which is equivalent to the leaky bucket.
- The implementation of Fig. 5.14.3 can be discussed under two different operating conditions, namely :
  1. For packets of fixed size.
  2. For packets of variable size.

1. Fixed size packets :

- If the arriving packets are of fixed size (e.g. cells in ATM networks), then the process of Fig. 5.14.3 will allow the removal of a fixed number of packets from the queue corresponding to every tick of the clock.

2. Packets of variable size :

- If the packets at the input of the process are of different size, then the fixed output rate will not correspond to the number of packets leaving the process but it will correspond to the number of bits leaving the process.

Algorithm :

- The algorithm for variable length packets is as follows :
  1. Initialize a counter to a number "n" at the tick of the clock.
  2. If "n" is greater than the packet size, then send the packet and decrement the counter by the packet size.
  3. Repeat step 2 until "n" becomes smaller than the packet size.
  4. Reset the counter and go back to step 1.

Note : Thus a leaky bucket algorithm shapes the bursty traffic to convert it into a fixed rate traffic. It does so by averaging the data rate. It drops the packets if the bucket (buffer) is full.

5.14.4 Token Bucket Algorithm :

MU : Dec. 10, May 15, Dec. 17

University Questions

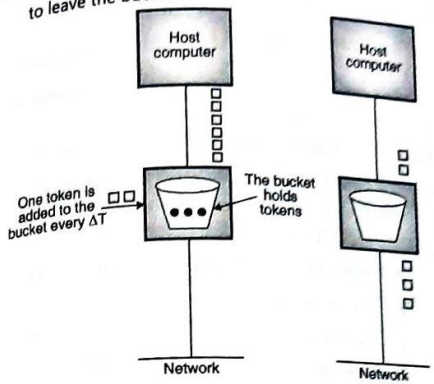
Q. 1 How does the token bucket algorithm works ? **(Dec. 10, 5 Marks)**

Q. 2 How does the Token Bucket Algorithm works ? **(May 15, 4 Marks)**

Q. 3 How does the token Bucket algorithm work ? **(Dec. 17, 5 Marks)**

- This algorithm is similar to the leaky bucket but it is possible to vary output rates.
- This is useful when larger burst of traffic is received.
- It enforces a long-term average transmission rate while permitting bounded bursts.
- In this approach, a token bucket is used to which manages the queue regulator that ultimately controls the rate of packet flow into the network.
- A token generator continuously produces tokens at a rate of R tokens per second and puts them into a token bucket with a depth of D tokens as shown in Fig. 5.14.4.
- If the token bucket gets full then the extra tokens are discarded.
- Token bucket algorithm is a variant of leaky bucket algorithm discussed earlier. Here the bucket is filled with tokens.

- A packet which grabs and destroys a token is allowed to leave the bucket.

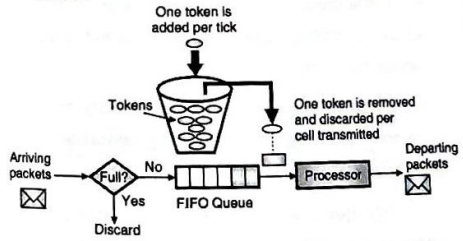


(G-483) **Fig. 5.14.4 : Token bucket algorithm**

- Due to this mechanism, the packets never get lost but they just have to wait to grab a token.
- At the same time, an unregulated stream of packets arrive and are placed into a packet queue that has a maximum length of L.
- If the flow delivers more packets than the queue can store, the excess packets are discarded.

Implementation of token bucket :

- Fig. 5.14.5 shows the implementation of token bucket.



(G-484) **Fig. 5.14.5 : Implementation of token bucket**

- The token bucket can be easily implemented with a counter. The token is initialized to zero.
- Every time a token is added, the counter is incremented by 1 and every time a packet is dispatched, the counter is decremented by 1.
- If the counter contents go to zero, the host cannot send any data.

Note : The token bucket allows the bursty traffic at maximum possible rate.

Token bucket performance :

Let, s = Burst length (seconds),
c = Bucket capacity (bytes),
ρ = Token arrival rate (bytes/second),
and m = Maximum source rate (bytes/second)

What is the duration of a maximum-rate burst through a token bucket ?

1. Maximum bytes sent from the token bucket during a burst is, $c + \rho \cdot s$
2. Maximum bytes the source can send during a burst is,
   $m \cdot s$
3. Setting the two equal and solving for s,
   $$s = \frac{c}{m - \rho}$$

5.14.5 Comparison of Token Bucket and Leaky Bucket :

MU : Dec. 15, Dec. 18, Dec. 19

University Questions

Q. 1 What is traffic shaping ? Explain leaky bucket algorithm and compare it with token bucket algorithm. **(Dec. 15, Dec. 18, 10 Marks)**

Q. 2 What is traffic shaping ? Explain leaky bucket algorithm and compare it with token Bucket algorithm. **(Dec. 19, 10 Marks)**

Table 5.14.1 : Comparison of Token Bucket and Leaky Bucket algorithm

Sr. No.	Parameter	Leaky Bucket	Token Bucket
1.	Principle of operation	Smooth out traffic by passing packets only when there is a token.	Smooths out traffic
2.	Permission to burstiness	No	Yes

Sr. No.	Parameter	Leaky Bucket	Token Bucket
3.	Discarding packet / token	Discards packet for which no token is available	Discards token when bucket is full, but never discards packets
4.	Queueing	No	Yes
5.	Application	Traffic shaping or traffic policing	Network traffic shaping or rate limiting

5.14.6 Combination of Token Bucket and Leaky Bucket :

- The token bucket and leaky bucket techniques can be combined to obtain the following advantages :
  1. To credit an idle host
  2. To regulate the traffic
- The token bucket is used first followed by the leaky bucket technique.
- The rate of leaky bucket needs to be higher than the rate of tokens dropped in the bucket.

5.14.7 Resource Reservation :

- The data flow is dependent on the following resources :
  1. Buffer
  2. Bandwidth
  3. CPU time
- The QoS can be improved by reserving these resources.
- The QoS model called integrated services operates on the principle of resource reservation, for improvement in QoS.

5.14.8 Admission Control :

- Admission control technique is used by a router or a switch.
- They use this mechanism to accept or reject the data flow based on predefined parameters called flow specifications.
- Before accepting a flow for processing, a router checks the flow specifications and finds out if it is possible to take up and handle this new data flow.
- It does this by comparing its bandwidth, buffer size, CPU speed etc. with the flow specifications.

5.15 Fragmentation :

- The network designers are not free to choose any size of the packet.

Factors deciding the size of packets :

- The maximum packet size varies network to network and the factors which decide the maximum packet size are as follows :
  1. Width of the TDM transmission slot.
  2. Protocols used.
  3. Type of operating system.
  4. International standards.
  5. Efforts to reduce retransmission.
  6. Desire to prevent one packet from occupying the channel too long.
- All these factors put a limit on the maximum packet size.
- The maximum payload size ranges from 48 bytes for an ATM cell to 65, 515 bytes for an IP packet.
- When a large packet wants to travel over a network whose maximum packet size is very small, we face a problem.
- The solution to this problem is to avoid this situation in the first place by using a routing algorithm which will avoid sending packets through the networks that cannot handle them.
- But this solution cannot be exercised every time. The real solution to this problem is **Fragmentation.**

Fragmentation :

- The technique in which, the gateways break up large packets into smaller ones called as fragments.
- After fragmentation of a large packet, each fragment is sent as a separate internet packet.
- But the reverse process of putting the fragments together is substantially difficult.

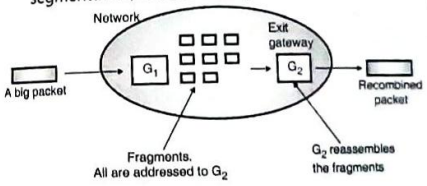
Recombination of fragments :

- The recombination of fragments can be done by using one of the following two strategies.

1. Transparent strategy and
2. Non-transparent strategy

5.15.1 Transparent Strategy :

- In this strategy, the fragmentation caused by a "small packet" network is made transparent to any subsequent network through which the packets will pass.
- When a large packet arrives at a gateway, G_1 in Fig. 5.15.1(a) it breaks the packet into fragments.
- Each fragment is then addressed to the same exit gateway. The exit gateway (G_2) recombines all these fragments.
- This strategy is illustrated in Fig. 5.15.1(a). In this way the small packet network has been made transparent i.e. the rest of the network cannot see what happened.
- The subsequent networks are not even aware that fragmentation has taken place.
- Fragmentation in ATM networks is called segmentation, but the concept is same.



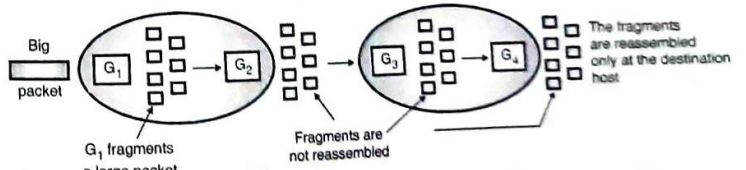
(G-494) Fig. 5.15.1(a) : Transparent strategy

Disadvantages :

- The disadvantages of transparent fragmentation are :
  1. The first problem with transparent fragmentation is that the exit gateway G_2 has to know that it has received all the pieces. For this a count field or an end of packet bit has to be included in each packet.
  2. Another important factor is that all the packets should exit via the same gateway.
  3. The last problem is the overhead required to repeatedly fragment and reassemble a large packet.

5.15.2 Non-transparent Strategy :

- In this strategy, the fragmented packets are not reassembled at any intermediate stage.
- That means the exit gateways will not reassemble the fragments.
- Instead each fragment is treated as a separate original packet.
- All these packets are passed through the exit gateway or gateways and their recombination is carried out at the destination host as shown in Fig. 5.15.1(b).
- This is called as a non-transparent fragmentation.



(G-495) Fig. 5.15.1(b) : Non-transparent strategy

Disadvantages :

- The disadvantages of non-transparent fragmentation are :
  1. Every host must be capable of reassembling the fragments.
  2. The total overhead increases due to fragmentation since each fragment has to have a header.
  3. When a packet is fragmented, the fragments will have to be numbered in such a way that the original data stream can be reconstructed at the destination.

Advantage :

- The advantage of non-transparent strategy is that now we can use multiple exit gateways and improve the network performance.

Review Questions

Q. 1 State and explain the network layer duties.
Q. 2 What are the network layer design issues ?
Q. 3 Explain the store and forward packet switching.
Q. 4 Write short notes on : implementing a connectioless service.
Q. 5 Compare virtual circuit and datagram subnets.
Q. 6 Explain the concept of unicast routing.
Q. 7 Explain the concept of unicast routing.
Q. 8 Explain the concept of broadcast routing.
Q. 9 Compare unicast and broadcat routing.
Q. 10 Explain the optimality principle.
Q. 11 Differentiate between static and dynamic routing.
Q. 12 What is flooding ?
Q. 13 Explain the concept of flow based routing.
Q. 14 Write short notes on : Hierarchical routing.
Q. 15 Write short notes on : distance vector routing.
Q. 16 Explain the looping problem in distance vector routing ?
Q. 17 What is multicast routing ?
Q. 18 Write short note on : Network layer congestion.
Q. 19 State the difference between multiple unicasting and multicasting.
Q. 20 Explain the concept of multicast address.
Q. 21 What is the difference between static and dynamic routing algorithms ?
Q. 22 Explain distance vector routing algorithm.
Q. 23 Write a short note on : Fragmentation.
Q. 24 Write a short note on : Count to infinity problem.
Q. 25 Explain the link state routing algorithm.
Q. 26 Write a short note on : Token Bucket Algorithm.
Q. 27 Explain duties performed by network layer.
Q. 28 Write a short note on : Leaky Bucket Algorithm.
Q. 29 Compare Leaky Bucket and Token Bucket Algorithms.
Q. 30 What is Quality of Service ?

□□□

Module 4

Chapter **6**

Network Layer Protocols

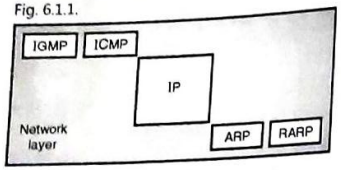
Syllabus

IPv4 Addressing (Classfull and Classless), Subnetting, Supernetting design problems, IPv4 Protocol. Network Address Translation (NAT), IPv6.
Protocols : ARP,RARP, ICMP, IGMP.

Chapter Contents

6.1 Network Layer Protocols :

- The main protocols corresponding to the network layer in the TCP/IP suite as well as Internet layer are : ARP, RARP, IP, ICMP and IGMP. This is as shown in Fig. 6.1.1.



(G-524)**Fig. 6.1.1 : Protocols at network layer**

- Out of these protocols IP is the most important protocol.
- It is responsible for host to host delivery of datagrams from a source to destination.
- But IP needs to take services of other protocols.
- IP takes help from ARP in order to find the MAC (physical) address of the next hop.
- IP uses the services of ICMP during the delivery of the datagram packets to handle unusual situations such as presence of an error.
- IP is basically designed for unicast delivery.
- But some new Internet applications as well as multimedia need multicast delivery.
- So for multicasting, IP has to use the services of another protocol called IGMP.
- IPv4 is the current version of IP whereas IPv6 is the latest version of IP.

6.1.1 Why IP Address ?

- How does the Internet Protocol (IP) know about the source of a datagram and its destination ?
- For a common user the Internet should appear as a single network and all the incompatibilities of the physical networks that make the Internet should remain hidden from the common user.
- Also the people connected to these physical networks should be able use any technology of their choice.
- So we need to have a common interface which binds the end users of Internet and the people dealing with their own networks.
- To identify each computer connected to the Internet uniquely is a great challenge.
- Different networking technologies have different physical addressing mechanisms.
- A **physical address** is also known as the **hardware address** and there are three methods to assign the hardware address to a computer as follows :
  1. Static addresses
  2. Configurable addresses
  3. Dynamic addresses.

1. Static addresses :

- The static address is a physical address which is hard coded in the Network Interface Card (NIC) of the computer.
- This address is provided by the network hardware manufacturer and it does not change ever.

2. Configurable addresses :

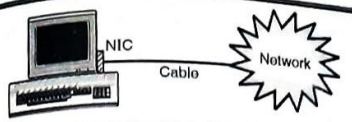
- In this method, the physical address is configured inside a computer at the time of its first installation at its site.
- The configurable address allows the user to set up a physical address.

3. Dynamic addresses :

- In this method, a server computer dynamically assigns a physical address to a computer every time it boots.
- Thus the physical address of a computer changes everytime it is switch off and on.
- **Note :** The method of static addresses is the simplest of all the three methods discussed so far.
- It is important to understand that every computer has a unique hardware or physical address and it is stored in the NIC of the computer.

Role of NIC :

- As discussed earlier, the NIC is an input/output interface on each computer.
- It allows the computer to communicate with all other computers on the network.
- This is as shown in Fig. 6.1.2.

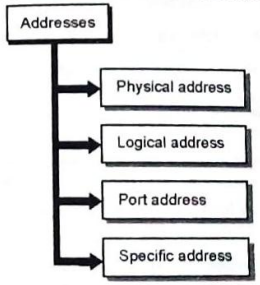


(G-1439) **Fig. 6.1.2 : Role of NIC**

- The NIC acts as an interface between a computer and its network.

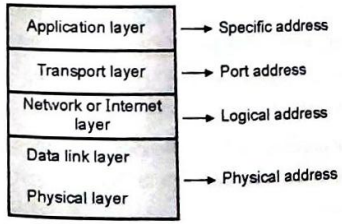
6.2 Addressing :

- When the computers wish to communicate with one another, they need to know the address of each other. Each computer has its own address.
- The addresses can be of different types such as physical addresses or logical address.
- In an internet employing the TCP/IP protocols, four levels of addresses are used by the computers.
  1. Physical address
  2. Logical address (IP)
  3. Port address and
  4. Specific address
- Fig. 6.2.1 shows the classification of addresses.



(G-75) **Fig. 6.2.1 : Classification of addresses in TCP/IP**

- Each of these addresses is associated with a specific layer of TCP/IP architecture as demonstrated in Fig. 6.2.2.



(G-76) **Fig. 6.2.2 : Relation between TCP/IP structure and addresses**

6.2.1 MAC Address (Physical Address) :

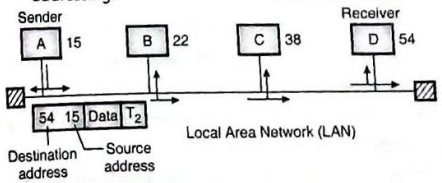
MU : May 08, May 09, Dec. 15

University Questions

Q. 1 Explain with example MAC address. **(May 08, May 09, 10 Marks)**

Q. 2 Differentiate between an IP address and a MAC or physical address. What is the need to map IP address to MAC address ? Explain which protocol does this similarly give a protocol which does reverse mapping. **(Dec. 15, 5 Marks)**

- The packets from source to destination hosts pass through physical networks.
- At the physical level the IP address is not useful but the hosts and routers are recognized by their MAC addresses.
- A MAC address is a local address. It is unique locally but it is not unique universally.
- The IP and MAC address are two different identifiers and both of them are needed, because a physical network can have two different protocols at the network layer at the same time.
- Similarly a packet may pass through different physical networks.
- So to deliver a packet to a host or a router, we require two levels of addressing namely IP addressing and MAC addressing.
- Most importantly we should be able to map the IP address into a corresponding MAC address.
- The size and format of the physical address varies depending on the nature of network.
- The Ethernet (LAN) uses a 48-bit (6-byte) physical address which is imprinted on the network interfacing card (NIC).
- Refer Fig. 6.2.3 which explains the concept of physical addressing.



(G-77) **Fig. 6.2.3 : Physical addresses**

- The sender computer with a physical address of 15 wants to communicate with the receiver computer with a physical address 54.
- The frame sent by the sender consists of the destination address, sender's address, encapsulated data and a trailer (T_2) that contains the error control bit.
- When this frame travels over the bus topology, every computer receives it and tries to match it with its own physical address.
- If the destination address in the frame header does not match with the physical address it will simply drop the frame.
- At receiver computer (D), the destination address matches with its physical address (54). So the frame is accepted and decapsulation is carried out to recover the data.
- The example of a 48 bit or 6 byte physical address is as follows. It contains 12-hexadecimal digits.

08 : 63 : 4C : 81 : 08 : 1D

6.2.2 Logical Addresses (IP Addresses) :

- Logical addresses are required to facilitate universal communications in which different types of physical networks can be involved.
- The logical address is also called as the IP (internet protocol) address.
- The Internet consists of many physical networks interconnected via devices like routers.
- Internet is a packed switched network that means the data form the source computer is sent in the form of small packets carrying the destination address upon them.
- A packet starts from the source host, passes through many physical networks and finally reaches the destination host.
- At the network level, the hosts and routers are recognised by their **IP addresses** or logical addresses.
- An IP address is an internetwork address. It is a universally unique address.
- Every protocol involved in internetworking requires IP addresses.
- The logical address used in internet is currently a 32-bit address.
- The same IP address can never be used by more than one computer on the Internet.

6.2.3 Port Address :

- The modern computers are designed to run multiple processes on it simultaneously.
- The main objective of internet is the process to process communication.
- For this purpose it is necessary to lable or name the processes.
- Thus the processes need addresses. The lable assigned to a process is called as a port address. It is a 16 bit address.

6.2.4 Specific Addresses :

- Some applications have user friendly addresses. The examples of specific addresses are the e-mail addresses or the University Resource Locators (URL).

6.3 ARP (Address Resolution Protocol) :

MU : Dec. 04, Dec. 09, Dec. 15, Dec. 18, May 19

University Questions

Q. 1 ARP and RARP both map addresses from one space to another. In this respect they are similar. In what major way do they differ ? **(Dec. 04, 4 Marks)**

Q. 2 What is Address Resolution Protocol (ARP) ? **(Dec. 09, 5 Marks)**

Q. 3 Differentiate between an IP address and a MAC or physical address. What is the need to map IP address to MAC address ? Explain which protocol does this similarly give a protocol which does reverse mapping. **(Dec. 15, 5 Marks)**

Q. 4 Write a short note on write ARP / RARP **(Dec. 18, 5 Marks)**

Q. 5 Write short note on : ARP. **(May 19, 5 Marks)**

- ARP as defined in RFC 826 is Ethernet Address Resolution Protocol.
- ARP provides service to IP, which make us think that it is in the link layer TCP/IP model (or DLL of OSI model).
- But its messages are carried by DLL protocol and are not capsulated within IP datagrams.

- That is why it can be called as a network layer protocol as well.
- Thus ARP occupies an unusual place in TCP/IP suite.
- But the most important point is that ARP provides an essential service when TCP/IP is running on a LAN.
- An Internet consists of various types of networks and the connecting devices like routers.
- A packet starts from the source host, passes through many physical networks and finally reaches the destination host.
- At the network level, the hosts and routers are recognised by their IP addresses.

IP address :

- An IP address is an internetwork address. It is a universally unique address.
- Every protocol involved in internetworking requires IP addresses.

MAC address :

- The packets from source to destination hosts pass through physical networks.
- At the physical level the IP address is not useful but the hosts and routers are addressed by their MAC addresses.
- A MAC address is a local address. It is unique locally but it is not unique universally.
- The IP and MAC address are two different identifiers and both of them are needed, because a physical network can have two different protocols operating at the network layer at the same time.
- Similarly a packet may travel through different physical networks.
- So to deliver a packet to a host or a router, we require addressing to take place at two levels namely IP addressing and MAC addressing.
- Most importantly we should be able to map the IP address into a corresponding MAC address.

6.3.1 Mapping of IP Address into a MAC Address :

- We have seen the need of mapping an IP address into a MAC address.
- Such a mapping can be of two types :

1. Static mapping and 2. Dynamic mapping

1. Static mapping :

- In static mapping a table is created and stored in each machine.
- This table associates an IP address with a MAC address.
- If a machine knows the IP address of another machine then it can search for the corresponding MAC address in its table.
- The limitation of static mapping is that the MAC addresses can change.
- These changed MAC addresses must be updated periodically in the static mapping table.

2. Dynamic mapping :

- In dynamic mapping technique a protocol is used for finding the other address when one type of address is known.
- There are two protocols used for carrying out the dynamic mapping.
- They are :
  1. Address Resolution Protocol (ARP).
  2. Reverse Address Resolution Protocol (RARP)
- The ARP is used for mapping an IP address to a MAC address whereas the RARP is used for mapping a MAC address to an IP address.

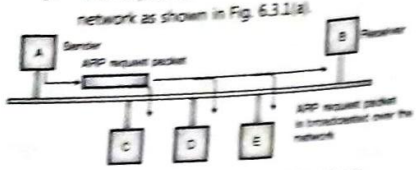
6.3.2 ARP Operation :

- ARP is used for mapping an IP address to its MAC address.
- For a LAN, each device has its own physical or station address as its identification.
- This address is stored on the NIC (Network Interface Card) of that machine.

How to find the MAC address ?

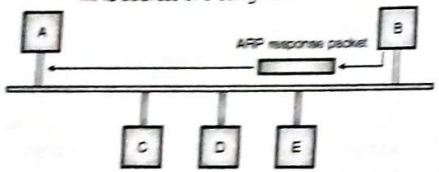
- When a router or a host (A) needs to find the MAC address of another host (B) the sequence of events taking place is as follows :
  1. The router or host A who wants to find the MAC address of some other router, sends an ARP request packet. This packet consists of IP and MAC addresses of the sender A and the IP address of the receiver (B).

2. This request packet is broadcasted over the network as shown in Fig. 6.3.1(a).



(G-575)**Fig. 6.3.1(a) : ARP request is broadcast**

3. Every host and router on the network will receive the ARP request packet and process it. But only the intended receiver (B) will recognize its IP address in the request packet and will send an ARP response packet back to A.
4. The ARP response packet has the IP and physical addresses of the receiver (B) in it. This packet is delivered only to A (unicast) using A's physical address in the ARP request packet. This is shown in Fig. 6.3.1(b). Thus host A has obtained the MAC address of B using ARP.



(G-576) **Fig. 6.3.1(b) : ARP response unicast**

6.3.3 Mapping Physical Address to Logical Address :

- Sometimes a host knows its physical address but needs to know its logical address.
- This can happen in the following two cases :
  1. If a diskless station has been just booted. This station can find its physical address by checking its interface but it does not know its logical address.
  2. An organization has less number of IP addresses. So it can not assign a separate IP address to each station. Hence it has to assign the IP addresses when a station demands for it.

6.3.4 ARP Cache Memory :

- The use of ARP would be inefficient if A needs to broadcast an ARP request for each IP packet that it to be sent to B, because instead of broadcasting the request it could have broadcast the IP packet itself.
- So ARP is efficient only if the ARP reply is stored in cache memory (cached) for a while.
- This is due to the fact that a system generally sends hundreds of packets to the same destination.
- Thus the system that receives an ARP reply stores the mapping in the cache memory and keeps it for 20 to 30 minutes.
- So if packets are again sent to the same destination then it could use this mapping instead of broadcasting an ARP request.
- Before sending an ARP request, the system checks its cache to see if the mapping could be found.

6.3.5 ARP Packet Format :

- The ARP message format is as shown in Fig. 6.3.2.

Hardware Type (16 bits)		Protocol type (16 bits)
Hardware length	Protocol length	Operation request 1, Reply 2
Sender hardware address		
Sender protocol address		
Target hardware address		
Target protocol address		

Fig. 6.3.2 : ARP message format

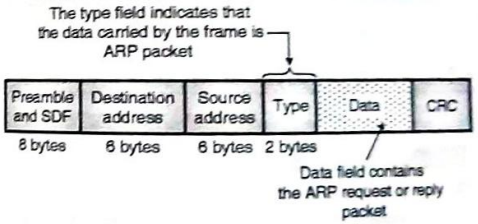
- The various fields in it are as follows :

1. **HTYPE (Hardware Type) :** This 16 bit field defines the type of network on which ARP is being run. ARP is capable of running on any physical network.
2. **PTYPE (Protocol Type) :** This 16 bit field is used to define the protocol using ARP. Note that we can use ARP with any higher-level protocol such as IPv4.
3. **HLEN (Hardware length) :** It is an 8 bit field which is used for defining the length of the physical address in bytes. For example, this value is 6 for Ethernet.
4. **PLEN (Protocol Length) :** This field is 8 bit long and it defines the length of the IP address in bytes. For IPv4 this value is 4.

5. **OPER (Operation) :** It is a 16 bit field which defines the type of packet. The two possible types of packets are : ARP request (1) and ARP reply (2).
6. **SHA (Sender Hardware Address) :** This field is used for defining the physical address of the sender. The length of this field is variable.
7. **SPA (Sender Protocol Address) :** This field defines the logical address of the sender. The length of this field is variable.
8. **THA (Target Hardware Address) :** It defines the physical address of the target. It is a variable length field. This field contains all zeros for the ARP request packet, because the receivers physical address is not known to the sender.
9. **TPA (Target Protocol Address) :** This field defines the logical address of the target. It is a variable length field.

6.3.6 Encapsulation :

- An ARP packet (request or reply) is inserted directly into the data link frame. Such an insertion is known as encapsulation.
- Fig. 6.3.3 shows an example of encapsulation in which an ARP packet being encapsulated in an Ethernet frame.



(G-578)**Fig. 6.3.3 : Encapsulation of ARP packet**

- The type field shows that the data carried by the frame is an ARP request or reply packet.

6.3.7 Operation of ARP on Internet :

- The services of ARP can be used under the following working conditions when it is being operated on Internet :
  1. The sender is a host and wants to communicate with another host which is on the same network.
  2. The sender is a host and wants to communicate with a host on another network.
  3. The sender is a router. It has received a datagram with a destination address of a host on another network.
  4. The sender is a router. It has received a datagram which is meant for a host in the same network.
- Now let us see how ARP works on the internet.

Operation :

1. The sender (host or router) knows the IP address of the target.
2. IP orders ARP to create an ARP request message. The request packet consists of senders physical and IP addresses plus the IP address of the target but the physical address of the target is not known.
3. This ARP request packet is sent to the data link layer. Here the ARP request packet is inserted in a frame.
4. Every router or host receives this frame because it is broadcast. All the machines except the target drop this packet as discussed earlier.
5. The target machine sends back a reply packet which contains the target's physical address. This reply is unicast and addressed only to the sender.
6. The sender receives the reply packet. Hence the physical address of the target has been obtained.
7. The IP datagram carrying data for the target machine is inserted in a frame and the frame is unicast to the target machine.

6.3.8 Four Different Cases :

- The four different cases in which the services of ARP can be used are as follows :

Case 1 :

- The host sender wants to send a packet to another host on the same network.
- In this case, the destination IP address in the datagram header acts as the logical address which should be mapped to a physical address.

Case 2 :

- This case corresponds to a situation where a host wants to send a packet to another host on another network.
- Here the host refers its routing table and finds the IP address of the next hop (router) for the destination host.
- If it does not have the routing table, then it will search for the IP address of the default router.
- The IP address of the router will be considered as the logical address which is to be mapped to the corresponding physical address.

Case 3 :

- In this case a router has received a datagram which is to be sent to a host on another network.
- To do this the router checks its routing table and finds the IP address of the next router.
- The IP address of the next router should be mapped to a physical address by the ARP.

Case 4 :

- The sender is a router. It has received a datagram which is to be sent to a host on the same network.
- In this case the IP address of the destination host should be mapped into a physical address.

6.4 The Reverse Address Resolution (RARP) Protocol :

MU : Dec. 04, Dec. 16, Dec. 18

University Questions

Q. 1 ARP and RARP both map addresses from one space to another. In this respect they are similar. In what major way do they differ ? **(Dec. 04, 4 Marks)**

Q. 2 Write short notes on : RARP **(Dec. 16, 5 Marks)**

Q. 3 Write a short note on the ARP / RARP. **(Dec. 18, 5 Marks)**

- ARP is used for solving the problem of finding out which Ethernet address corresponds to a given IP address.
- That means ARP is used for the mapping of IP address to physical or MAC address.
- But sometimes we have to solve a reverse problem.
- That means we have to obtain the IP address corresponding to the given Ethernet (MAC) address.
- Such a problem can occur when booting a diskless workstation.
- The problem of obtaining the IP address when an Ethernet address is given, can be solved by using RARP (Reverse Address Resolution Protocol).
- The newly booted workstation is allowed to broadcast its Ethernet address.
- The RARP server after receiving this request, checks the Ethernet address in its files and finds the corresponding IP address. This IP address is then sent back.
- The disadvantage of RARP is that it uses a destination address of all 1s (limited broadcasting) to reach the RARP server.
- But such broadcasts are not forwarded by routers, so a RARP server is needed on each network.
- In order to get around this problem, another bootstrap protocol called BOOTP has been invented.
- Unlike RARP, it uses UDP messages which are forwarded over routers.
- It also provides a diskless workstation with additional information, including the IP address of the file server holding the memory image, the IP address of the default router and the subnet mask to use.

6.5 Internet Protocol Version 4 (IPv4) :

MU : May 15, May 16, Dec. 17

University Questions

Q. 1 What is the function of IP protocol ? Discuss its header format. **(May 15, 10 Marks)**

Q. 2 What is IPv4 protocol ? Explain the IPv4 header format with diagram. **(May 16, Dec. 17, 10 Marks)**

- We have already discussed the addressing mechanism, for the IP packets.
- Now we will discuss the format of IP packet in the next few sections.
- In the discussion we will see that an IP packet consists of a base header and options which are sometimes useful in controlling the packet delivery.

Relation between IP and other protocols :

- The main protocols corresponding to the network layer in the TCP/IP suite as well as Internet layer are : ARP, RARP, IP, ICMP and IGMP.
- Out of these protocols IP is the most important protocol.
- It is responsible for host to host delivery of datagrams from a source to destination.
- But IP needs to take services of other protocols.
- IP takes help from ARP in order to find the MAC (physical) address of the next hop.
- IP uses the services of ICMP during the delivery of the datagram packets to handle unusual situations such as presence of an error.
- IP is basically designed for unicast delivery.
- But some new Internet applications as well as multimedia need multicast delivery.
- So for multicasting, IP has to use the services of another protocol called IGMP.
- IPv4 is the current version of IP whereas IPv6 is the latest version of IP.

6.5.1 Internet Protocol (IP) :

MU : May 11, Dec. 11, Dec. 18

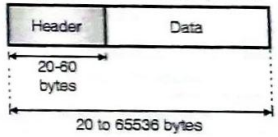
University Questions

Q. 1 What is IPv4 protocol ? Explain the IPv4 header format with diagram. **(May 11, Dec. 11, Dec. 18, 10 Marks)**

- The Internet Protocol is the host to host delivery protocol which belongs to the network layer and is designed for the Internet.
- IP is used as the transmission mechanism by the TCP / IP protocols.
- That means the TCP or UDP packets are encapsulated in the IP packet and the IP carries it from source to destination.
- IP is a connectionless datagram protocol with no guarantee of reliability.
- It is an unreliable protocol because it does not provide any error control or flow control.
- IP can only detect the error and discards the packet if it is corrupted.
- If IP is to be made more reliable, then it must be paired with a reliable protocol such as TCP at the transport layer.
- Each IP datagram is handled independently and each one can follow a different route to the destination.
- So there is a possibility of receiving out of order packets at the destination. Some packets may even be lost or corrupted.
- IP relies on a higher level protocol to take care of all these problems.
- The version of IP that we are going to discuss is called as IPv4 i.e. IP version 4.
- IP is also called as a best effort delivery protocol.
- The meaning of the term best effort delivery is that the IP packet can get lost or corrupted or delayed.
- They may arrive out of order at the destination or may create congestion in the network.

6.5.2 Datagrams (IP Packet) :

- Packets in IP layer are called datagrams. Fig. 6.5.1 shows the typical format of an IP packet.



(G-525) **Fig. 6.5.1 : IPv4 datagram format**

- A datagram has two parts namely the header and data as shown.
- The length of datagram is not fixed. It varies from 20 bytes to 65536 bytes.
- The length of the header is 20 to 60 bytes.
- The information necessary for the routing and delivery of the datagram has been stored in the header.
- The other part of the datagram is the data field which is of variable length.
- It is a custom in TCP/IP to show the header in 4-byte (32 bit) sections.

6.5.3 IPv4 Header Format :

MU : May 10, May 11, Dec. 11, May 12, May 13, Dec. 13, May 15, May 16, Dec. 17, Dec. 18, May 19

University Questions

Q. 1 Draw and explain the structure of IP Frame Header. **(May 10, 10 Marks)**

Q. 2 What is IPv4 protocol ? Explain the IPv4 header format with diagram. **(May 11, Dec. 11, 10 Marks)**

Q. 3 Describe the IPv4 header format in detail. **(May 12, May 13, 10 Marks)**

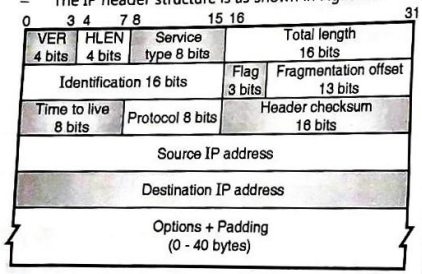
Q. 4 Write short notes on : IP header format. **(Dec. 13, 10 Marks)**

Q. 5 What is the function of IP protocol ? Discuss its header format. **(May 15, 10 Marks)**

Q. 6 What is IPv4 protocol ? Explain the IPv4 header format with diagram. **(May 16, Dec. 17, Dec. 18, 10 Marks)**

Q. 7 Describe IPv4 header format with diagram. **(May 19, 10 Marks)**

- The IP frame header contains routing information and control information associated with datagram delivery.
- The IP header structure is as shown in Fig. 6.5.2.



(G-2082) Fig. 6.5.2 : IPv4 header format

- Various fields in the header format are as follows :

1. VER (Version) :

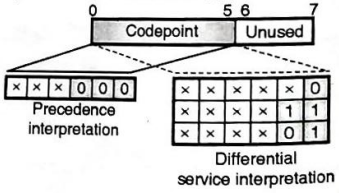
- This is a 4 bit field which is used to define the version of IP protocol.
- The current version of IP is 4 i.e. IPv4 but in future it may be completely replaced by the latest version of IP i.e. IPv6.
- This field will indicate the IP software running on the processing machine that this datagram belongs to IPv4 version.
- If the processing machine is using some other version of IP, then the datagram will be discarded.

2. HLEN (Header length) :

- This 4-bit long field is used for defining the length of the datagram header in 4-byte words.
- The value of this field is multiplied by 4 to get the length of the IPv4 header which varies between 20 and 60 bytes.
- When there are no options, the value of this field is 5 and the **header length** is 5 × 4 = 20 bytes.
- When the value of option field is maximum the value of HLEN field is 15 and the corresponding header length is maximum i.e. 15 × 4 = 60 bytes.

3. Service type :

- In the earlier designs of IP header, this field was called as **Type of Service (TOS)** field and its job was to define how the datagram should be handled.
- At that time, a part of this field used to define the precedence of datagram and the remaining part used to define the type of service out of different possible services such as low delay, high throughput etc.
- But now the interpretation of this field has been changed by IETF.
- This field is now supposed to define a set of **differential services.** Fig. 6.5.3 illustrates the new interpretation of the **service type** field.



(G-2083) Fig. 6.5.3 : New interpretation of service type field

- As seen in Fig. 6.5.3, in the new interpretation, the service type field is divided into two subfields namely, the 6 bit **codepoint** subfield and a 2 bit **unused** subfield.
- We can use the 6-bit **codepoint** subfield in two different ways, as follows :

1. For the purpose of precedence interpretation.

2. For the differential service interpretation.

- If the three right most bits are zeros, then the three leftmost bits are interpreted the same as the precedence bits in the service field (old interpretation).
- That means it is compatible with the old interpretation of this field.
- The precedence interpretation is used for defining the priority level of this datagram (from 0 to 7) in the situations like congestion.
- In the event of congestion, the datagrams with lowest precedence (0) will be discarded first.

Differential service interpretation :

- When the three rightmost bits are not all zeros, the 6 bit codepoint subfield is used for differential service interpretation.
- In that case these 6 bits can be used for defining a total of 56 (64 – 8) services, on the basis of the priorities assigned by the Internet or local authorities as per Table 6.5.1.

Table 6.5.1 : Values of codepoints

Category	Codepoint	Assigning authority
1.	× × × × × 0	Internet
2.	× × × × 1 1	Local
3.	× × × × 0 1	Temporary or Experimental

- The first, second and third categories contain 24, 16 and 16 service types respectively.
- The Internet authorities assign the first category.
- The local authorities assign the second while the third one is temporary and can be used for experimental purposes.

4. Total length :

- This 16 bit field is used to define the total length of the IP datagram.
- The total length includes the length of header as well as the data field.
- The field length of this fields is 16 bits so the total length of the IP datagram is restricted to $(2^{16} - 1)$ = 65535 bytes out of which 20 to 60 bytes constitute the header and the remaining bytes are reserved to carry data from upper layers.
- This field allows the length of a datagram to be upto 65,535 bytes, although such long datagrams are impractical for most hosts and networks.
- All hosts must be prepared to accept datagram of upto 576 bytes, regardless of whether they arrive whole or in the form of fragments.
- The hosts are recommended to send datagram larger than 576 bytes only if the destination is prepared to accept larger datagram.
- We can find the length of data by subtracting the header length from the total length.
- As stated earlier the header length can be obtained by multiplying the contents of HLEN field by four.

∴ Length of data = Total length – header length

- The total length (maximum value) of 65,535 bytes might seem to be large but in future the size of IP datagram is likely to increase further because the improvement in technology will allow more bandwidth.

Why do we need the total length field ?

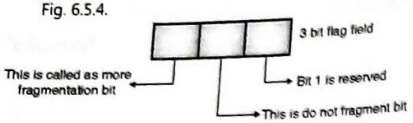
- We might feel that the **total length** field is not at all required because the host or router will drop the header and trailer when it receives a frame.
- Then why to include this field ?
- The answer to this question is that in many situations we do not need this field at all.
- But in some special situations, only the datagram is not encapsulated in the frame but there are some padding bits as well that are included.
- In such situations, the machine (host or router) that decapsulates the datagram, needs to check the **total length** field so as to understand how much is the data and how much is the padding ?

5. Identification :

- This field is used to identify the datagram originating from the source host.
- When a datagram is fragmented, the contents of the identification field get copied into all fragments.
- This identification number is used by the destination to reassemble the fragments of the datagram.

6. Flags :

- This is a three bit field. The 3 bits are as shown in Fig. 6.5.4.

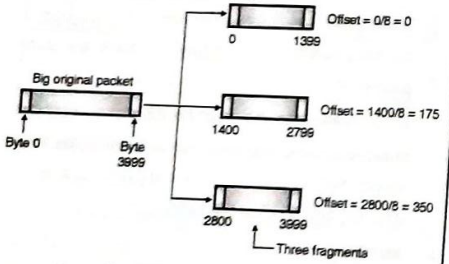


(G-527)Fig. 6.5.4 : Flag bits

- First bit is reserved, and it should be 0.
- The second bit is known as the "Do Not Fragment" bit.
- If this bit is "1" then machine understands that the datagram is not to be fragmented.
- But if the value of this bit is 0 then the machine should fragment the datagram if and only if necessary.
- The third bit is known as "More Fragment Bit" (M). M = 1 indicates that the datagram is not the last fragment and M = 0 indicates that this is the last or the only fragment.

7. Fragmentation offset :

- This is a 13 bit field which is used to indicate the relative position of this fragment with respect to the complete datagram.
- It is the offset of the data in the original datagram measured in units of 8 bytes.
- To understand this refer Fig. 6.5.5.



(G-528)Fig. 6.5.5 : Example of fragmentation

- The original IP packet (datagram) contains 4000 bytes numbered from 0 to 3999. It is fragmented into three fragments.
- The first fragment contains 1400 bytes numbered from 0 to 1399.
- The offset for this fragment is 0/8 = 0. Similarly the offsets for the other two fragments are 1400/8 = 175 and 2800/8 = 350 respectively as shown in Fig. 6.5.5.
- The offset is measured in units of 8 bytes. Because the length of the offset field is 13 bits, so the fragments should be of size such that first byte number is divisible by 8.

8. Time to Live (TTL) :

- This is an 8-bit field which controls the maximum number of routers visited by the datagram during its lifetime.
- A datagram has a limited lifetime for travelling through an Internet.
- Originally the TTL field was designed to hold the **timestamp.**
- This timestamp value was decremented by one, everytime the datagram visits a router.
- As soon as the timestamp value reduces to zero the datagram is discarded.
- But for this scheme to become successful, all the machines must have synchronized clocks and they must know the time taken by a datagram to travel from one router to the other.
- Today the TTL field is used to **control** the maximum number of hops i.e. router by a datagram.
- At the time of sending a datagram, the source host will store a number in the TTL field.
- This number is approximately twice the maximum number of routers present between any two hosts.
- Everytime this datagram visits a router, this value is decremented by one.
- If after decrementing, the value of TTL field reduces to zero then that router discards the datagram.

Need of TTL field :

- Sometimes the routing tables in the Internet get corrupted, due to which a datagram may travel between two or more routers for a very long time but never ever gets delivered to the destination host.
- The TTL field is needed in such situations for **limiting the lifetime of a datagram.**

- The TTL field is also used to limit the journey of a packet intentionally.
- For example if a packet is to be confined to a local network only then a 1 is stored in the TTL field of this packet.
- As soon as it reaches the first router, then TTL field value is decremented from 1 to 0 and the packet will be discarded.

9. Protocol :

- This is an 8-bit field which is used for defining the higher level protocol which uses the services of IP layer.
- The data from different high level protocols can be encapsulated into an IP datagram.
- These protocols could be UDP, TCP, ICMP, IGMP etc.
- The protocol field contents would tell the name of the protocol at the final destination to which this IP datagram is to be delivered.
- At the destination, the value of this field helps in the process of demultiplexing.
- Table 6.5.2 shows some of the values of this field corresponding to different high level protocols.

Table 6.5.2

Value	Protocol	Value	Protocol
1	ICMP	17	UDP
2	IGMP	89	OSPF
6	TCP		

10. Header checksum :

- A checksum in IP packet covers on the header only.
- Since some header fields change, this field is recomputed and verified at each point that the Internet header is processed.

11. Source address :

- This field is used for defining the IP address of the source. It is a 32 bit field.

12. Destination address :

- This field is used for defining the IP address of the destination. It is also a 32 bit field.

13. Options :

- Options are not required for every datagram.
- They are used for network testing and debugging.
- We have discussed all the options in detail, later in this chapter.

Ex. 6.5.1 : Explain IPv4 header format in detail. If value at HLEN field is 1101 find the size of option and padding field ?

Dec. 15, 10 Marks

Ans. :

- Please refer Section 6.5.3 for IPv4 header format.

Total length of the IP header = HLEN contents × 4 bytes

= 13 × 4 = 52 bytes.

- Number of bytes corresponding to all the fields except the option + padding is 20 bytes.

∴ Size of option + padding field = 52 – 20 = 32 bytes

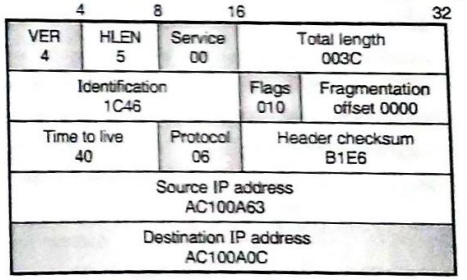
...Ans.

Ex. 6.5.2 : A IP header from an IP packet received at destination 4500003c1c4640004006b16ac100a63 ac100a0c. Map these values to IP header and explain all bits.

Dec. 14, 10 Marks

Soln. :

- The IP packet data received at the destination is in the form of hexadecimal digits (4 bits each).
- The IP header is as shown in Fig. P. 6.5.2(a). In the received IP packet data there are 40 hex digits (4 bits each) or 20 bytes.



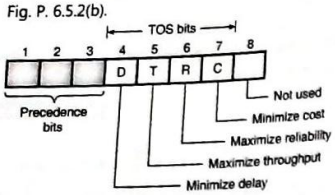
(G-1780) Fig. P. 6.5.2(a) : IPv4 header format

- The received data has been placed in various fields of the IP header as shown in Fig. P. 6.5.2(a).

Explanation :

1. **Version (VER)** = 4. Therefore the version is IPv4.
2. **Header length (HLEN)** = 5. Therefore the header length is 5 × 4 = 20 bytes.

3. **Service** = 00. This field indicates differentiated services. These 8 bits are bifurcated as shown in Fig. P. 6.5.2(b).



(G-1781) **Fig. P. 6.5.2(b) : Service type or differentiated services**

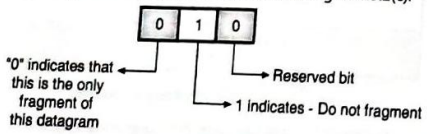
- From Fig. P. 6.5.2(b),
- **Precedence bits** are 000, which shows that this packet has the least priority in the event of congestion.
- **TOS bits** are 0000 which indicates that the type of service is normal service or default service.

4. **Total length** = 003C. This defines the total length of the IPv4 datagram (header + data) in bytes.

Total length = 003C
= 36 bytes (header + data)
∴ Number of data bytes = 36 – 20 (header bytes)
= 16 bytes.

5. **Identification :** This is a 16 bit field. Its value here is **1C46**. It is the identification number of this datagram. If it is fragmented, then this number will be copied into all the fragments. At the destination the same number will be used to reassemble all the fragments of this datagram.

6. **Flags :** The contents are as shown in Fig. P. 6.5.2(c).



(G-1782) **Fig. P. 6.5.2(c) : Flag bits**

7. **Fragmentation offset :**
- *This is a 13 bit field and they are all zeros in this case.*
- *So the number of the first byte obtained by multiplying offset value by 8 will be 0.*

8. ***Time To Live :*** *TOL = 40 in this case.*
- This value is approximately twice the maximum number of routers between any two hosts.
- When this datagram passes through a router it will decrement this number by 1. When TOL becomes 00 the router will discard this datagram.

9. **Protocol** = 06. That means the higher level protocol which uses the services of the IP layer is TCP.

10. **Checksum** = B1E6 This is used at the destination to verify if the header has been received without any errors.

11. **Source IP address** = AC100A63 it is a 32 bit address.

12. **Destination IP address** = AC100A0C it is also a 32 bit address.

6.6 IPv4 Addresses :

- Each computer connected to the Internet should be identified uniquely.
- The identifier used for this purpose is called as the **Internet address** or IP address.
- The hosts and routers on the Internet have unique IP addresses.
- The current version of IP (Internet Protocol) is IPv4 whereas the advanced version is IPv6.
- The IPv4 address is a 32-bit address and it is used for defining the connection of a host or router to the Internet.
- Thus an IP address is an address of the interface.

6.6.1 Uniqueness of IP Addresses :

- The IP address is **unique** and **universal.** That means each IP address defines only **one connection** to the Internet.
- At any given time, no two devices connected to the Internet can have the same IP address.
- But if a device is connected to the Internet via two connections through two different networks, then it can have two different IP addresses.
- All the IPv4 addresses are 32 bit long and they are used in the source address and destination address fields of the IP header.
- The IP addresses for hosts are assigned by the network administrator.

- For Internet it has to be obtained from the network information center.

6.6.2 Address Space :

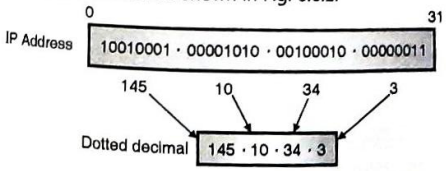
- The IPv4 protocol has an address space. It is defined as the total number of addresses used by the protocol.
- If N number of bits are used for defining an address then the address space will be 2^N addresses.
- For IPv4, N is 32 bits. Hence its address space is 2^{32} or 4, 294, 967, 296 (more than 4 billion).
- So theoretically more than 4 billion devices could be connected to the Internet.
- Thus **the address space** of IPv4 is 2^{32}.

6.6.3 Notation :

- The IPv4 addresses can be shown use three different notations as follows :
  1. Binary notations (base 2).
  2. Dotted decimal notation (base 256).
  3. Hexadecimal notation (base 16).
- Out of these the **dotted decimal** notation is most commonly used.

Dotted decimal notation :

- This notation has become popular because of the two advantages it offers.
- This notation makes the IPv4 address more compact and easy to read.
- The 32 bit IPv4 address is grouped into groups of 8-bits each separated by decimal points (dots).
- Each 8-bit group is then converted into an equivalent decimal number as shown in Fig. 6.6.1.

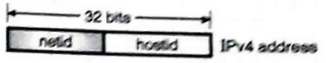


(G-2001) **Fig. 6.6.1 : Dotted decimal notation**

- Each octet (byte) can take a value between 0 and 255.
- Therefore the IPv4 address in the dotted decimal notation has a range from 0.0.0.0 to 255.255.255.255.
- For example the IPv4 address of 1001 0001.00001010 00100010 00000011 is denoted in the dotted decimal form as 145.10.34.3.

6.6.4 IPv4 Address Format :

- A 32 bit IPv4 address consists of two parts. The first part is called as **net id** i.e. network identification which identifies a network on the Internet and the second part is called as the **host id** which identifies a host on that network.
- Fig. 6.6.2 shows the IPv4 address format. Note that the **net id** and **host id** are of variable lengths depending on the class of address.
- Note that class D and E addresses are not divided into net id and host id for the reasons discussed later on.



(G-2002) **Fig. 6.6.2 : IPv4 address format**

6.7 Classful Addressing : MU : May 10

University Questions

Q. 1 Explain classful addressing in IPv4. **(May 10, 10 Marks)**

- The concept of IP addresses is few decades old. It uses the concept of **classes.**
- This architecture is called as the **classful addressing.**
- Later on in mid 1990s a new architecture of addressing was introduced which was known as **classless addressing.**
- This new architecture has superseded the original architecture.
- In this section we are going to discuss the classful addressing.

6.7.1 IPv4 Address Classes :

MU : Dec. 05, Dec. 06, Dec. 07, May 08, May 09, May 10, May 17, Dec. 18

University Questions

Q. 1 Describe the classification of IP-addresses in IPv4. **(Dec. 05, 5 Marks)**

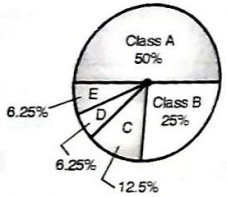
Q. 2 Explain the following term : IP address. **(Dec. 06, Dec. 07, May 08, May 09, 10 Marks)**

Q. 3 Discuss various special IP addresses. **(May 09, 10 Marks)**

Q. 4 Explain classful addressing in IPv4. **(May 10, 10 Marks)**

Q. 5 Explain with examples the classification of IPv4 addresses. **(May 17, 5 Marks, Dec. 18, 4 Marks)**

- In the classful addressing architecture, the IP address space has been divided into five classes : A, B, C, D and E.
- Fig. 6.7.1 shows the percentage of occupation of the address space by each class.
- The number of class A addresses is the highest i.e. 50% and those of classes D and E is the lowest i.e. 6.25%. (G-530)



Class	No. of addresses	
A	2^{31}	50%
B	2^{30}	25%
C	2^{29}	12.5%
D	2^{28}	6.25%
E	2^{28}	6.25%

(G-2003) **Fig. 6.7.1 : Classful addressing occupation of address space**

6.7.2 Formats of Various Classes :

MU : Dec. 05, Dec. 06, Dec. 07, May 08, May 09, May 17

University Questions

Q. 1 Describe the classification of IP-addresses in IPv4. **(Dec. 05, 5 Marks)**

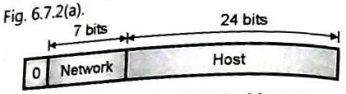
Q. 2 Explain the following term : IP address. **(Dec. 06, Dec. 07, May 08, May 09, 10 Marks)**

Q. 3 Discuss various special IP addresses. **(May 09, 10 Marks)**

Q. 4 Explain with examples the classification of IPv4 addresses. **(May 17, 5 Marks)**

Class A format :

- The formats used for IPv4 address are as shown in Fig. 6.7.2.
- The IPv4 address for class A networks is shown in Fig. 6.7.2(a).

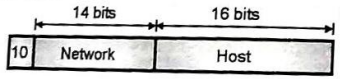


(G-531) **Fig. 6.7.2(a) : Class A IPv4 address formats**

- The network field is 7 bit long as shown in Fig. 6.7.2(a) and the host field is of 24 bit length.
- So the network field can have numbers between 1 to 126.
- But the host numbers will range from 0.0.0.0 to 127.255.255.255.
- Thus in class A, there can be 126 types of networks and 17 million hosts.
- The "0" in the first field identifies that it is a class A network address.

Class B format :

- The class B address format is shown in Fig. 6.7.2(b).
- The first two fields identify the network, and the number in the first field must be in the range 128 - 191.

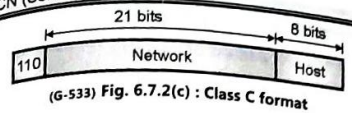


(G-532) **Fig. 6.7.2(b) : Class B format**

- Class B networks are large. Host numbers 0.0 and 255.255 are reserved, so there can be upto 65,534 (216-2) hosts in a class B network.
- Most of the 16,382 class B addresses have been allocated.
- The first block covers address from 128.0.0.0 to 128.255.255.255 and the last block covers from 191.255.0.0 to 191.255.255.255.
- **Example :** 128.89.0.26, for host 0.26 on net 128.89.

Class C format :

- The class C address format is shown in Fig. 6.7.2(c).

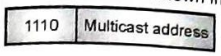


(G-533) **Fig. 6.7.2(c) : Class C format**

- The first block in class C covers addresses from 192.0.0.0 to 192.0.0.255 and the last block covers addresses from 223.255.255.0 to 223.255.255.255.

Class D format :

- The class D address format is shown in Fig. 6.7.2(d).

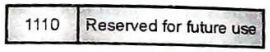


(G-2847) **Fig. 6.7.2(d) : Class D format**

- The class format allows for upto 2 million networks with upto 254 hosts each and class D format allows the multicast in which a datagram is directed to multiple hosts.

Class E address format :

- Fig. 6.7.2(e) shows the address format for a class E address.
- This address begins with 11110 which shows that it is reserved for the future use.

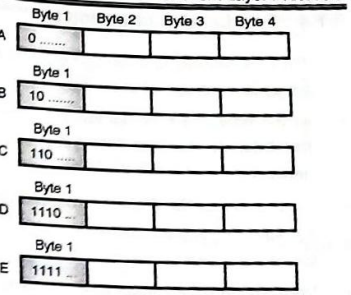


(G-2848) **Fig. 6.7.2(e) : IPv4 address for class E network**

- The 32 bit (4 byte) network addresses are usually written in dotted decimal notation.
- In this notation each of the 4-bytes is written in decimal from 0 to 255.
- So the lowest IP address is 0.0.0.0 i.e. all the 32 bits are zero and the highest IPv4 address is 255.255.255.255.

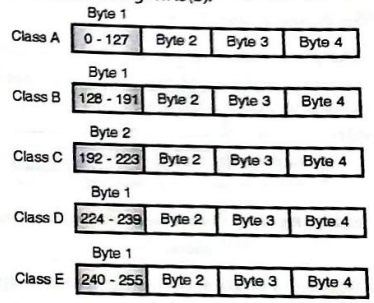
6.7.3 How to Recognize Classes ?

- When an IPv4 address is given to us either in the binary or dotted decimal notation, we can find the class of the address.
- If the given address is in the binary notation then we can identify its class by inspecting the first few bits of the address.
- This is as shown in Fig. 6.7.3(a).



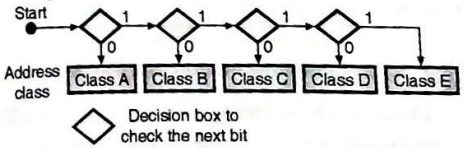
(G-2004) **Fig. 6.7.3(a) : Finding the address class**

- If the given address is in the dotted decimal notation then we can identify the address class by inspecting the first byte of the address.
- This is as shown in Fig. 6.7.3(b).



(G-2005) **Fig. 6.7.3(b) : Finding the address class**

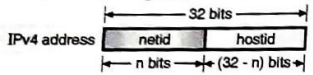
- It is important to note here that there are some special addresses which fall in class A or E.
- These special addresses are to be treated as the exceptions to the classful addressing.
- In computers, the IPv4 addresses are generally stored in the binary notation format.
- Therefore it is possible to write an algorithm which can identify the address class by using the continuous checking process.
- The principle of such an algorithm has been shown in Fig. 6.7.4.



(G-2006) **Fig. 6.7.4 : Algorithm to identify address class**

6.7.4 Two Level Addressing :

- The IPv4 addressing is used for defining a destination for an Internet packet at the network layer.
- At the time when classful addresses were designed, the Internet was considered as the network of networks.
- In other words the whole Internet was divided into a number of smaller networks with many hosts connected to each network.
- Normally an organization which wants to connect to the Internet creates a network and the Internet authorities allocate a block of address to the organization.
- These addresses can be in class A, B or C.
- All the addresses allotted to an organization belong to a single block.
- Therefore each IPv4 address in classful addressing system is made up of two parts namely **net id** and **host id** as shown in Fig. 6.7.5.



(G-2007) Fig. 6.7.5 : Two level addressing in classful addressing

- The job of the **net id** is to define a network and that of the **host id** is to define a particular host in that network.
- As shown in Fig. 6.7.5 if n bits define **net id** then the remaining (32-n) bits define **host id.**
- The value of "n" is not same for all the classes. Infact it is depend on the class as shown in Table 6.7.1.

Table 6.7.1

Class	Value of n
A	n = 8
B	n = 16
C	n = 24

6.7.5 Extracting Information in a Block :

- A block is nothing but a range of addresses. For any given block we would be interested to extract the following three pieces of information :
  1. The total number of addresses in the block.
  2. The first address of the block.
  3. The last address in the block.
- Before extracting all this information, we have to identify the class of the address as discussed earlier.
- Once we find the class of the block, we will have the values of "n" (the length of **net id** in bits) and (32 – n) i.e. the length of the **host id** in bits.
- It is now possible to obtain the three pieces of information mentioned above as shown in Fig. 6.7.6.

1. Total number of addresses in the block :

- The total number of IPv4 addresses in the given block will be equal to,

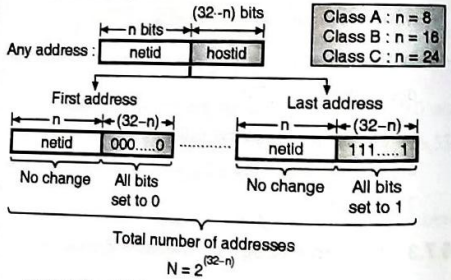
$$N = 2^{(32-n)} \qquad ...(6.7.1)$$

2. First address in the block :

- The first address in the given block can be obtained by keeping the leftmost "n" bits in the address as it is and setting all the (32 – n) rightmost bits to 0 as shown in Fig. 6.7.6.

3. Last address in the block :

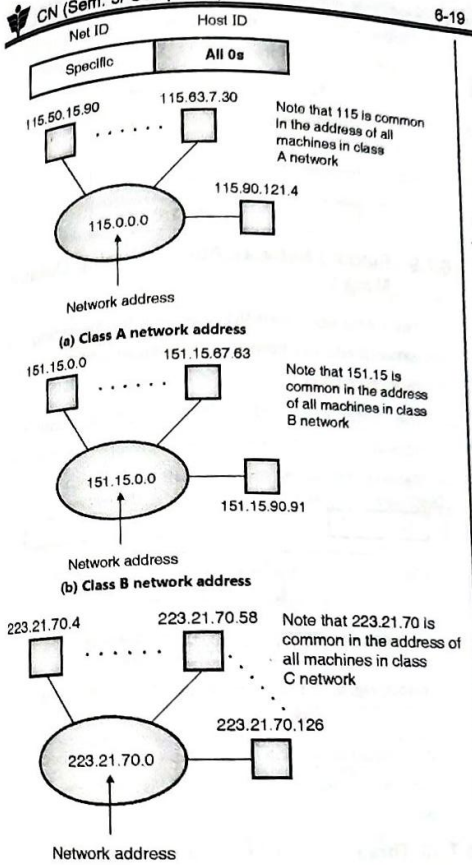
- The last address in the given block can be obtained by keeping the leftmost "n" bits in the address as it is and then setting all the (32 – n) rightmost bits to 1 as shown in Fig. 6.7.6.



(G-2008) Fig. 6.7.6 : Information extraction in classful addressing

6.7.6 Network Address :

- The network address is an address that defines the network itself.
- It cannot be assigned to a host. Fig. 6.7.7 shows the examples of network addresses for different classes.



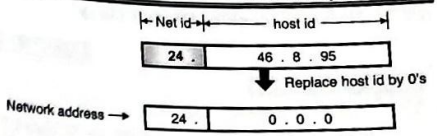
(G-536) Fig. 6.7.7

- The following examples will enable you to find the network address.

Ex. 6.7.1 : For the address 24.46.8.95 identify the type of network and find the network address.

Soln. :

- Examine the first byte. Its value is 24 i.e. it is between 0 and 127.
- So it is a class A network.
- So only the first byte defines the Net id. So we can find the network address by replacing the host id with 0s.
- The process of obtaining the network address is shown in Fig. P. 6.7.1.



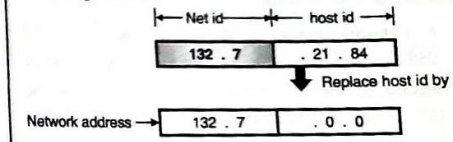
(G-537) Fig. P. 6.7.1

- So the network address is 24.0.0.0.

Ex. 6.7.2 : For the address 132.7.21.84 find the type of network and the network address.

Soln. :

- Examine the first byte. It is 132 i.e. between 128 and 192.
- So it is a class B network.
- So the first two bytes define the net id. Replace the host id with 0's to get the network address as shown in Fig. P. 6.7.2.



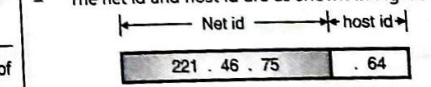
(G-538) Fig. P. 6.7.2

- So the network address is 132.7.0.0.

Ex. 6.7.3 : Find the class of the network if the address is 221.46.75.64.

Soln. :

- The first byte is 221 i.e. between 192 and 255. So this is a class C network.
- The net id and host id are as shown in Fig. P. 6.7.3.



(G-539) Fig. P. 6.7.3

What is the difference between net id and network address ?

- The network address is different from a net id. A network address has both net id and host id, with 0s for the host id.

Where to use the network address ?

- The network address is used to route the packets to the desired location.

6.7.7 Network Mask or Default Mask :

MU : Dec. 06

University Questions

Q. 1 Explain subnetting and masking with suitable examples. **(Dec. 06, 10 Marks)**

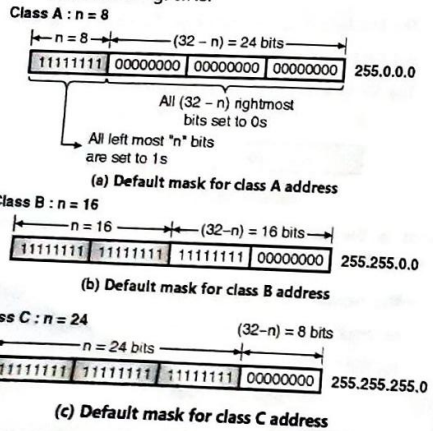
- Earlier we have discussed the methods for extracting different pieces of information.
- But all these methods are theoretical methods which are useful in explaining the concept.
- But practically these methods are not used. When a packet arrives at the input of the router in the Internet, it uses an algorithm to extract the **network address** from the destination address in the received packet.
- This can be achieved by using a **network mask.**

Definition of default mask :

- A **network mask** or **default mask** in classful addressing is defined as a 32-bit number obtained by setting all the "n" leftmost bits to 1s and all the (32 – n) rightmost bits to 0.

6.7.8 Default Masks for Different Classes :

- We know that the value of n is different for different classes. Therefore their default masks also will be different.
- The default masks for class A, B and C addresses are as shown in Fig. 6.7.8.



(G-2009) Fig. 6.7.8

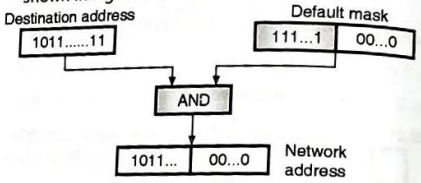
- Table 6.7.2 enlists the default masks of the three classes of IPv4 addresses.

Table 6.7.2 : Default masks

Address class	Default mask
A	255.0.0.0
B	255.255.0.0
C	255.255.255.0

6.7.9 Finding Network Address using Default Mask :

- The router uses the AND operation for extracting the network address from the destination address of the received packet.
- The router ANDs the destination address with the default mask to extract the network address as shown in Fig. 6.7.9.



(G-2010) Fig. 6.7.9 : Finding a network address using the default mask

- It is possible to use the default mask to find the number of addresses and the last address in the block.

6.7.10 Three Level Addressing : Subnetting :

MU : May 08, Dec. 08, May 16, Dec. 18, Dec. 19

University Questions

Q. 1 Explain subnetting and supernetting. **(May 08, Dec. 08, 10 Marks)**

Q. 2 Explain in short subnetting. **(May 16, 4 Marks)**

Q. 3 Explain the need of subnet mask in subnetting. **(Dec. 18, 4 Marks)**

Q. 4 What is subnetting ? What are the default subnet masks ? **(Dec. 19, 5 Marks)**

- As discussed earlier, the originally designed IP addresses were with two level addressing with **net id** and **host id.**
- The two level addressing is based on the principle that in order to reach a host on the Internet, we have to reach the network first and then the host.

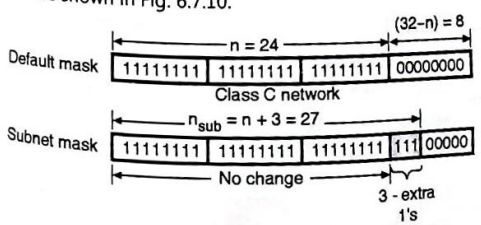
- But very soon it became evident that the two level addressing would not be sufficient for the following two reasons :
  1. First it was needed to divide a large network of an organization (to which a block in class A or B is allotted) into many smaller **subnets** (subnetworks) for improved management and security.
  2. Second reason is more important. The blocks in class A and B were almost depleted and the blocks in class C were smaller than the needs of most organization. Therefore the organizations had to divide their allotted class A or B block into smaller subnetworks and share them.

Definition of subnetting :

- We can define the **subnetting** as the principle of splitting a block of addresses into smaller blocks of addresses.
- In the process of **subnetting** we divide a big network into smaller subnetworks or **subnets.**
- Each such subnet has its own **subnet address.**

Subnet mask :

- The **network mask** or **default** mask that we discussed earlier is used when the given network is **not** to be divided into smaller subnetworks i.e. when **subnetting** is **not** to be done.
- But when the given network is to be divided into smaller subnets i.e. when subnetting is to be done, we need to create a **subnet mask** for each subnet.
- Fig. 6.7.10 shows the format of a subnet mask. Each subnet has its own **net id** and **host id.**
- If we want to divide a network into 8 subnets then the corresponding subnet mask will have three extra 1's because $2^3 = 8$, as compared to the default mask, as shown in Fig. 6.7.10.



(G-2011) Fig. 6.7.10 : Default and subnet masks

- In Fig. 6.7.10, we have shown the default mask and subnet mask when a class C network is to be divided into 8 subnets.

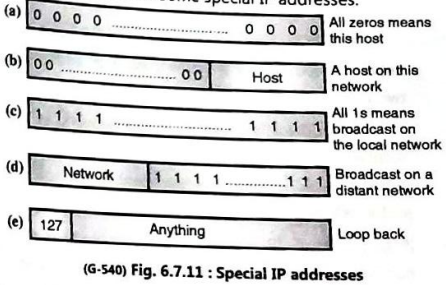
6.7.11 Special IP Addresses :

MU : May 09

University Questions

Q. 1 Discuss various special IP addresses. **(May 09, 10 Marks)**

- Fig. 6.7.11 shows some special IP addresses.



(G-540) Fig. 6.7.11 : Special IP addresses

- All zeros means this host or this network and all 1s means broadcast address to all hosts on the indicated network.
- The IP address 0.0.0.0 is used by the hosts when they are being booted but not used afterward.
- The IP addresses with 0 as the network number refer to their own network without knowing its number as shown in Fig. 6.7.11(b).
- The address having all ones is used for broadcasting on the local network such as a LAN as shown in Fig. 6.7.11(c).
- Refer Fig. 6.7.11(d). This is an address with proper network number and all 1s in the host field.
- This address allow machines to send broadcast packets to distant LANs anywhere in the Internet.
- If the address is "127. Anything" as shown in Fig. 6.7.11(e) then it is a reserved address **loopback testing.**
- This feature is also used for debugging network software.

6.7.12 Limitations of IPv4 :

- The most obvious limitation of IPv4 is its address field. IP relies on network layer addresses to identify end-points on networks, and each networked device has a unique IP address.

- IPv4 uses a 32-bit addressing scheme, which gives it 4 billion possible addresses.
- With the proliferation of networked devices including PCs, cell phones, wireless devices, etc., unique IP addresses are becoming scarce, and the world could theoretically run out of IP addresses.
- If a network has slightly more number of hosts than a particular class, then it needs either two IP addresses of that class or the next class of IP address.
- For example, let use say a network has 300 hosts, this network needs either a single class B IP address or two class C IP addresses.
- If class B address is allocated to this network, as the number of hosts that can be defined in a class B network is $(2^{16} - 2)$, a large number of host IP addresses are wasted.
- If two class C IP addresses are allocated, as the number of networks that can be defined using a class C address is only (2^{21}), the number of available class C networks will quickly exhaust.
- Because of the above two reasons, a lot of IP addresses are wasted and also the available IP address space is rapidly reduced.
- Other identified limitations of the IPv4 protocol are : Complex host and router configuration, non-hierarchical addressing, difficulty in re-numbering addresses, large routing tables, non-trivial implementations in providing security, QoS (Quality of Service), mobility and multi-homing, multicasting etc.
- To overcome these problems the internet protocol version 6 (IPv6) which is also known as internet protocol, next generation (IPng) was proposed.
- In IPv6 the internet protocol was extensively modified for accommodating the unforeseen growth of the internet.
- The format and length of the IP addresses has been changed and the packet format also is changed.

6.7.13 Classless Addressing :

- Eventhough the number of actual devices connected to Internet is much less than 4 billion, the address depletion has taken place due to flaws in the classful addressing scheme.
- We have run out of class A and B addresses. To overcome these problems, the classless addressing is now being tried out.
- In the classless addressing, there are no classes but the address generation take place in blocks.

Address blocks :

- Address block is defined as the range of addresses.
- In the classless addressing, when an entity wants to get connected to the internet, a block (range) of addresses is granted to it.
- The size of this block i.e. number of addresses depends on the size of the entity as well as its nature.
- That means for a small entity such as a household only one or two addresses will be given whereas for a larger entity like an organization, thousands of addresses can be allotted.

Restrictions :

- Some of the restriction on classless address blocks have been imposed by the internet authorities in order to simplify the process of address handling.
  1. The addresses in a block should be continuous, i.e. serial in manner.
  2. The total number of addresses in a block has to be equal to some power of 2 i.e. 2^1, 2^2, 2^3 ...etc.
  3. The first address should be evenly divisible by the number of addresses.

6.7.14 Supernetting :

- The class A and class B addresses are almost depleted. But class C addresses are still available.
- But the size of class C address with a maximum number of 256 addresses does not satisfy the needs of an organization. More addresses will be required.
- The solution to this problem is **supernetting.**
- In supernetting an organization combines several class C blocks to create a large range of addresses i.e. several networks are combined to create a supernetwork.
- By doing this the organization can apply for a set of class C blocks instead of just one.

Example of supernetting :

- If an organization needs 1000 addresses, they can be obtained by using four C blocks (one C block corresponds to 256 addresses).
- The organization can then use these addresses as one supernetwork as a whole.

Note : The classful addressing is almost obsolete now and it is being replaced with classless addressing.

6.7.15 Who Decides the IP Addresses ?

- No two IP addresses should be same. This is ensured by a central authority that issues the prefix or the network number portion of the IP address.
- Locally an ISP is to be contacted in order to get a unique IP address prefix.
- At the global level the Internet Assigned Number Authority (IANA) allots an IP address prefix to the ISP.
- Thus it is ensured that the IP addresses are not duplicated.
- Conceptually IANA is a wholesales and ISP is a retailer of the IP addresses because ISP purchases IP addresses from IANA and sells them to the customers.

6.7.16 Registered and Unregistered Addresses :

- Registered IP addresses are required for computers which are accessible from the Internet but not every computer that is connected to the Internet.
- For security reasons, networks use firewalls or some other technologies for protecting the computers.
- The firewalls will enable the workstations to access the Internet but do not allow the other systems on the Internet to access them.
- These workstations are given the unregistered private IP addresses.
- These addresses are assigned by the network administrator without obtaining them from an ISP (Internet Service Provider) or IANA.
- These are special network addresses in each class as shown in Table 6.7.3.

Table 6.7.3 : IP addresses for private networks

Class	Network address
A	10.0.0.0 through 10.255.255.255
B	172.16.0.0 through 172.31.255.255
C	192.168.0.0 through 192.168.255.255

- These addresses are to be used for private networks and are called **unregistered addresses.**
- We can choose any of these unregistered address while building our own private network.

6.7.17 Solved Examples :

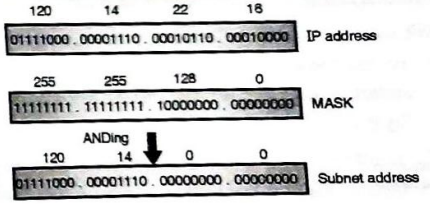
Ex. 6.7.4 : Find the sub-network address and the host id for the following :

Sr. No.	IP address	MASK
(a)	120.14.22.16	255.255.128.0
(b)	140.11.36.22	255.255.255.0
(c)	141.181.14.16	255.255.224.0
(d)	200.34.22.156	255.255.255.240

Soln. :

Step 1 : To find the subnet address :

- In order to find the subnet address we have to AND the IP address and the mask as follows :

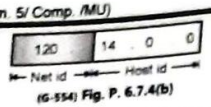


(G-553) Fig. P. 6.7.4(a)

- So the subnet address is 120.14.0.0.
- Similarly we can find the other subnet addresses.

Step 2 : Host Id :

- Examine the first byte of the subnet address. It is 120 which is between 0 and 127. Hence this is a class A network.
- So only the first byte corresponds to the net id and the remaining three bytes correspond to the host id as shown in Fig. P. 6.7.4(b).



(G-554) Fig. P. 6.7.4(b)

- So the host id is 14.0.0. Similarly we can find the other host id.

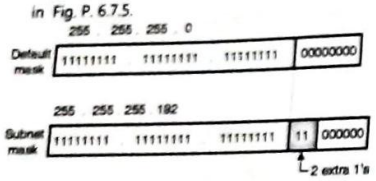
Ex. 6.7.5 : The IP address of a host on class C network is 198.123.46.237. Four networks are allowed for this network. What is subnet mask ?

Soln. :

- The default mask for a class C network is,

  255.255.255.0

- In order to have four networks, we must have two extra 1s.
- Hence the default mask and subnet mask are shown in Fig. P. 6.7.5.



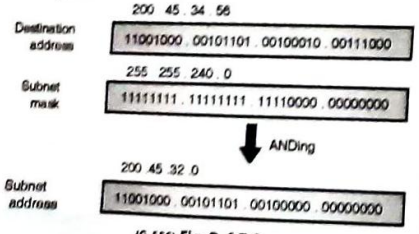
(G-555) Fig. P. 6.7.5

- Thus the required subnet mask is 255.255.255.192.

Ex. 6.7.6 : What is the subnet address if the destination address is 200.45.34.56 and subnet mask is 255.255.240.0 ?

Soln. :

- To find the subnet address we have to AND the IP address and the subnet mask as shown in Fig. P. 6.7.6.



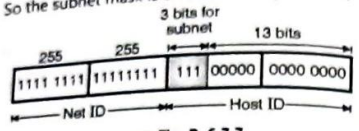
(G-556) Fig. P. 6.7.6

- Thus the required subnet address is 200.45.32.0.

Ex. 6.7.7 : Perform the subnetting of the following IP address 160.111. X.X Original subnet mask 255.255.0.0 Number of subnets 6 (six)

Soln. :

- The original subnet mask indicates that we are dealing with a class B address.
- In order to have six subnets we need to use 3 extra bits from the bits that are reserved for host ID.
- So the subnet mask is as shown in Fig. P. 6.7.7.



(G-566) Fig. P. 6.7.7

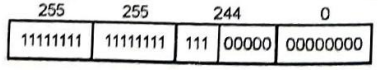
- The 3-bits reserved for subnetting will have 8 combinations from 000 to 111 out of which any six combinations can be used for 6 subnets.
- Let us decide that the combinations 000 to 001 are not to be used.
- Then the subnet masks for the 6 possible subnets will have the following addresses.

Subnet 1	255.255.64.0
Subnet 2	255.255.96.0
Subnet 3	255.255.128.0
Subnet 4	255.255.160.0
Subnet 5	255.255.192.0
Subnet 6	255.255.224.0

Ex. 6.7.8 : A class A network on the internet has a subnet mask of 255.255.224.0. What is the maximum number of hosts per subnet ? **May 05, 4 Marks**

Soln. :

- A subnet mask of 255.255.224.0 corresponds to the following pattern. **(G-2846)**

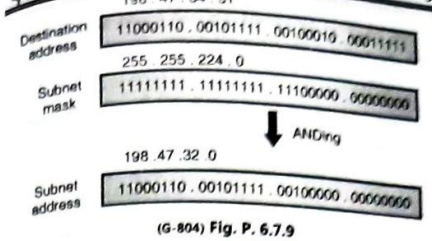


- Due to 3 additional 1s (shaded portion) there will be $2^3 = 8$ subnets and the number of hosts per subnet will be $2^{13} = 8192$.

Ex. 6.7.9 : What is subnet address if the destination address is 198.47.34.31 and subnet mask is 255.255.224.0 **Dec. 09, 5 Marks**

Soln. :

- To find subnet address we have to AND the IP address and the subnet mark as shown Fig. P. 6.7.9.



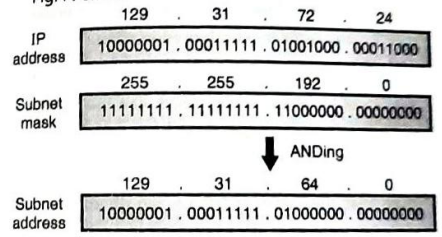
(G-804) Fig. P. 6.7.9

- Thus the required subnet address is 198.47.32.0

Ex. 6.7.10 : What is subnetting ? What are the default subnet masks ? Find the subnet address if the IP address is 129.31.72.24 and subnet mask is 255.255.192.0. **Dec. 15, 5 Marks**

Soln. :

- Please refer section 6.7.10 for subnetting.
- To find the subnet address we have to AND the IP address and the subnet mask as shown in Fig. P. 6.7.10.



(G-1868) Fig. P. 6.7.10

- Thus the required subnet address is 129.31.64.0.

Ex. 6.7.11 : What is subnetting ? Given the class C network 192.168.10.0 use the subnet mask 255.255.255.192 to create subnets and answer the following :

1. What is the number of subnets created ?
2. How many hosts per subnet ?
3. Calculate the IP address of the first host, the last host and the broadcast address of each subnet.

May 17, 10 Marks

Soln. :

- For subnetting refer section 6.7.10.

Given : IP address : 192.168.10.0 (class C)

Subnet mask : 255.255.255.192

Step 1 : Number of subnets and number of hosts :

255.255.255.192 ...(Given)

11111111 . 11111111 . 11111111 . 11000000

- The number of subnets are determined by the number of extra 1's.

∴ Number of extra 1's = 2

∴ Number of subnets = $2^2 = 4$ **...Ans.**

- The value of n is 26 which means the number of hosts per subnet is,

$2^{32-26} = 2^6 = 64$ **...Ans.**

Step 2 : IP address of the first host, last host and broadcast address :

- The following is the range of subnets :

Subnet	Subnet range
1.	192.168.10.0 to 192.168.10.63
2.	192.168.10.64 to 192.168.10.127
3.	192.168.10.128 to 192.168.10.191
4.	192.168.10.192 to 192.168.10.255

IP address of first host : 193.129.65.1

IP address of last host : 254.190.126.62

Broadcast address : 255.191.127.63

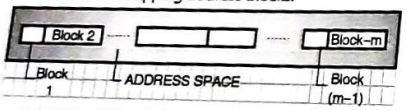
6.8 Classless Addressing in IPv4 :

- Eventhough the number of actual devices connected to Internet is much less than 4 billion, the address depletion has taken place due to flaws in the classful addressing scheme.
- We have run out of class A and B addresses. To overcome these problems, the super netting and subnetting has been tried as discussed earlier.
- But subnetting and supernetting also could not solve the problem of address depletion in IPv4.
- Due to increased number of Internet users, it was evident that a larger address space would be required as a long term solution to this problem.
- For this the length of the IP address should be increased which means the IP packet itself must be changed.
- A long term solution is to switch to IPv6.

- But a short term solution which uses the same address space has been deviced for IPv4. It is known as **classless addressing.**
- In the classless addressing, there are no classes but the address generation take place in blocks.
- The classless addressing was announced by the Internet authorities in 1996 in which blocks of variable length which do not belong to any class are used.

6.8.1 Variable Length Blocks :

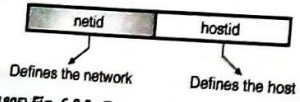
- Address block is defined as the range of addresses.
- In the classless addressing, when an entity wants to get connected to the internet, a block (range) of addresses is granted to it.
- The size of this block i.e. number of addresses depends on the size of the entity as well as its nature.
- That means for a small entity such as a household only one or two addresses will be given whereas for a larger entity like an organization, thousands of addresses can be allotted.
- Fig. 6.8.1 shows how the address space is divided into non overlapping address blocks.



(G-1804) **Fig. 6.8.1 : Variable length blocks in classless addressing**

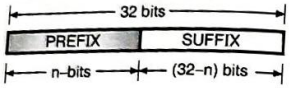
Two level addressing :

- We have discussed the two level addressing for classfull addressing which divided an address into two parts namely : net id and host id as shown in Fig. 6.8.2.



(G-1805) **Fig. 6.8.2 : Two layer addressing in classfull addressing**

- The **net id** and **host id** define the network and host respectively.
- It is possible to use the same idea in the classless addressing as well.
- A block of addresses granted to an organization is divided into two parts called as the **prefix** and the **suffix.**
- The role of prefix is same as that of the net id whereas as the role of suffix is same as that of the host id.
- Thus in a block granted to an organization, all the addresses will have the **same prefix** but each address will have a different **suffix.**
- Thus the prefix defines the network (organization to which the address block has been granted) while the suffix defines individual hosts on the network.
- The concept of two level addressing in classless addressing using the prefix and suffix is as shown in Fig. 6.8.3.



(G-1806) **Fig. 6.8.3 : Two level addressing using prefix and suffix for classless addressing**

- The IPv4 address is 32 bit long out of which the prefix will be of length "n" which can take any value from 0 to 32 and the length of the suffix will be (32 – n) bits.
- Note that the value of "n" i.e. length of the prefix depends on the length of the address block allotted (granted) to an organization.

Ex. 6.8.1 : Find out the values of prefix and suffix lengths in classless addressing if all the available addresses in IPv4 is to be considered as one single block.

Soln. :

- The total addresses in IPv4 is 2^{32} = 4,294,967,296.
- We have to consider this as one block hence the prefix length n = 0.
- Whereas all the hosts will have their individual addresses. So all the 32 bits will be allotted to the suffix length.

Ex. 6.8.2 : For the same data of the previous example find out the values of prefix and suffix lengths if all the available IPv4 addresses are divided into 4,294,967,296 blocks with each block having only one host.

Soln. :

- Here the prefix length for each block is n = 32, and the suffix length would be (32-n) = 0.
- The address of the single host in each block will be same as its block address itself.

Note : The two previous examples show that the prefix number n and the number of addresses in a block are inversely proportional to each other. With increase in the value of n, the number of addresses in a block will decrease.

6.8.2 The Slash Notation (CIDR Notation) :

MU : Dec. 09, May 11, May 13, May 16

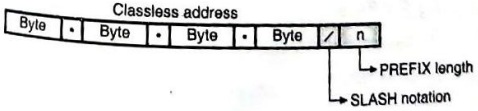
University Questions

Q. 1 Explain Classless Inter Domain Routing (CIDR) **(Dec. 09, May 11, 5 Marks)**

Q. 2 Write short notes on : CIDR. **(May 13, 5 Marks)**

Q. 3 Explain Classless Inter Domain Routing (CIDR). **(May 16, 10 Marks)**

- If an address (classful or classless) is given to us and we want to extract information from it, then the net id in classful addressing or the prefix in classless addressing are extremely important and useful to us.
- However it is not easy to identify the prefix bits in a given classless address. It is easy to identify the net id from the given classful address.
- For a given classless address it is not possible to find the prefix length because the given address can belong to a block with any prefix length.
- Therefore, in classless addressing it is essential to include the prefix length to each address if the block of the given address is to be found.
- Hence the prefix length "n" is added to the classless address separated by a **slash** and the notation is known as the **slash notation.**
- Fig. 6.8.4 demonstrates a classless address with slash notation.



(G-1807) **Fig. 6.8.4 : Slash notation**

- The slash notation is also called as Classless Interdomain Routing or CIDR notation.

6.8.3 Network Mask :

- We have discussed the concept of network mask in the classful addressing.
- The same concept is also applicable in the classless addressing as well.
- A **network mask** in classless addressing is a 32 bit number.
- With its "n" left most bits (corresponding to the prefix) all set to 1s and the remaining (32-n) bits corresponding to the suffix all set to 0s.

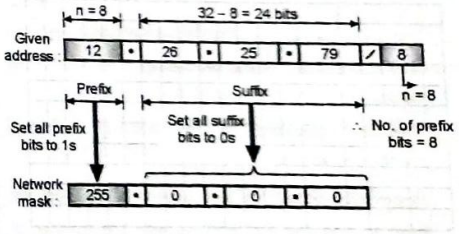
Ex. 6.8.3 : For the following addresses identify the number of prefix bits and write down the network mask :

1. 12.26.25.79 / 8
2. 130.12.230.156 / 16

Soln. :

1. Classless CIDR address : 12.26.25.79 / 8

- As per the slash notation we have n = 8 i.e. number of prefix bits is 8.
- Therefore the number of suffix bits = 32 – 8 = 24.
- In order to obtain the network mask the prefix bits all set to 1s and the suffix bits all set to zero as shown in Fig. P. 6.8.3(a).

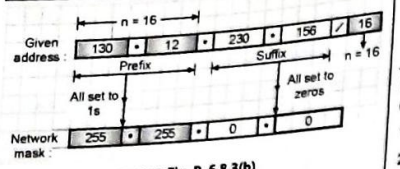


(G-1808) **Fig. P. 6.8.3(a)**

- Thus the network mask = 255.0.0.0

2. Classless CIDR Address : 130.12.230.156 / 16

- As per the slash notation, n = 16 i.e. number of prefix bits is 16.
- Number of suffix bits = 32 – 16 = 16
- In order to obtain the network mask, set all the prefix bits to 1s and set all the suffix bits to 0s as shown in Fig. P. 6.8.3(b).



(G-1809) Fig. P. 6.8.3(b)

- Thus the network mask = 255.255.0.0

6.8.4 Extracting the Block Information :

- We can extract all the required information from the given classless address in the CIDR notation. The information that we can obtain is as follows :
  1. The first address (network address)
  2. The number of addresses.
  3. The last address.
- We can obtain the number of addresses in a block as follows :

  Number of addresses in a block $N = 2^{(32-n)}$ ...(6.8.1)

  Where n = Number of prefix bits.
- The first address or network address in block can be obtained by ANDing the address with the network mask.

  First address = (Any address) AND (Network mask) ...(6.8.2)
- OR what we can do is keep the "n" leftmost bits of any address as it is and set the remaining (32-n) bits to 0s.
- This is equivalent to the ANDing operation mentioned above.
- In order to obtain the last address in the block we have to add the first address with the number of addresses in the block directly.

∴ Last address = First address + Number of addresses in the block ...(6.8.3)

- It is also possible to obtain the last address by ORing the address with complement of the network mask.

∴ Last address = (Any address) OR [NOT (Network Mask)] ...(6.8.4)

- One more way of obtaining the last address of the block is to keep all the "n" left most bits (prefix bits) as it is and set all the (32-n) bits (suffix bits) to 1s.

Ex. 6.8.4 : If an address in a block is given in CIDR classless notation as 64.32.16.8 / 27 then find the following :

1. Number of addresses in the block (N)
2. The first address and
3. The last address.

Soln. :

Step 1 : Find n :

Given address = 64.32.16.8 / 27

Hence n = 27 from the slash notation.

∴ n = 27 bits.

∴ Prefix bits = 27, suffix bits = 32 – 27 = 5

Step 2 : Number of addresses in the block (N) :

$N = 2^{(32-n)} = 2^5 = 32$

Step 3 : Find the first address :

- Refer Fig. P. 6.8.4(a) to obtain the first address in the block.
- For this we have to AND the given address with the network mask.

	n	(32 – n)
Network mask =	27 ones	5 zeros

∴ Network mask = 255.255.255.224

- For ANDing write the given address and network mask in their binary notations as shown in Fig. P. 6.8.4(a).

∴ From Fig. P. 6.8.4(a) we get the first address in the block as :

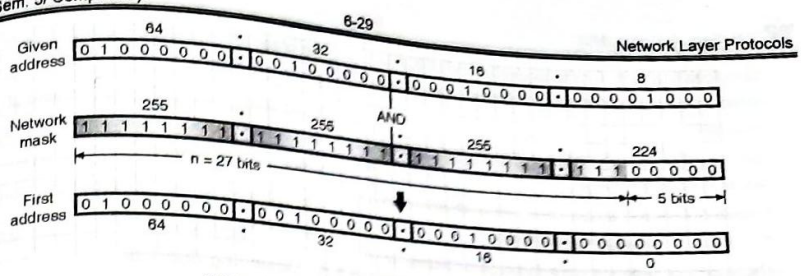
First address = 64.32.16.0 **...Ans.**

Step 4 : Find the last address :

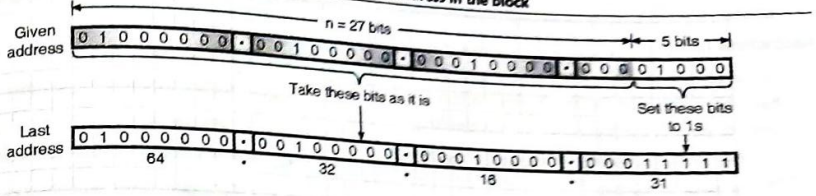
- To obtain the last address in the block, we have to keep the left most 27 bits in the given address as it is and set the remaining 5 bits to 1s as shown in Fig. P. 6.8.4(b).

∴ From Fig. P. 6.8.4(b) we get the last address in the block as follows :

Last address = 64.32.16.31



(G-1810) Fig. P. 6.8.4(a) : First address in the block



(G-1811) Fig. P. 6.8.4(b) : Last address

Ex. 6.8.5 : For the classless address 129.65.33.01 / 24 find the following :

1. Number of addresses in the block (N)
2. The first address.
3. The last address.

Soln. :

Step 1 : Find n :

- Given address = 129.65.33.01 / 24 hence n = 24 from the slash notation.

  ∴ n = 24 bits

∴ Prefix bits = 24, suffix bits = 32 – 24 = 8

Step 2 : Number of addresses in the block (N) :

$N = 2^{(32-n)} = 2^8 = 256$ **...Ans.**

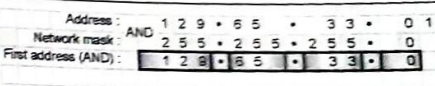
Step 3 : Find the first address :

- Refer Fig. P. 6.8.5(a) to obtain the first address in the block.
- For this we have to AND the given address with the network mask.

	n	(32 – n)
Network mask =	24 ones	8 zeros

∴ Network mask = 255.255.255.0

- For ANDing write the given address and network mask in their dotted decimal notations as shown.



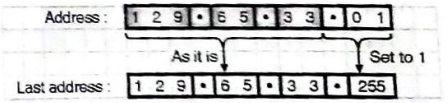
(G-1812) Fig. P. 6.8.5(a) : First address in the block

∴ From Fig. P. 6.8.5(a) we get the first address in the block as :

First address = 129.65.33.0 **...Ans.**

Step 4 : Find the last address :

- To obtain the last address in the block, we have to keep the left most 24 bits in the given address as it is and set the remaining 8 bits to 1s as shown in Fig. P. 6.8.5(b).



(G-1813) Fig. P. 6.8.5(b) : Last address in the block

- From Fig. P. 6.8.5(b) we get, the last address in the block is as follows :

Last address = 129.65.33.255 **...Ans.**

6.8.5 Block Allocation :

- Now let us understand how to allocate the blocks in the classless addressing.

- The global authority for the block allocation is ICANA means Internet Corporation for Assigned Names and Addresses.
- But the individual addresses of the Internet users is not allotted by the ICANA.
- Instead ICANA will assign large blocks of addresses to various ISPs or large organizations.
- These ISPs or organization will assign addresses to the individual Internet users from their allotted blocks.

Restrictions :

- Some of the restriction on classless address blocks have been imposed by the internet authorities in order to simplify the process of address handling.
  1. The addresses in a block should be continuous, i.e. serial in manner.
  2. The total number of addresses in a block has to be equal to some power of 2 i.e. $2^1, 2^2, 2^3$ ...etc.
  3. The first address should be evenly divisible by the number of addresses.

6.8.6 Relation to Classful Addressing :

- The classful addressing may be imagined as the special case of classless addressing such that the blocks of addresses in class A, B and C type addresses will have the prefix lengths $n_A = 8$, $n_B = 16$ and $n_C = 24$.
- Table 6.8.1 lists the prefix lengths for class A to F classful addresses and using this information we can change a block in classful addressing to a block in classless addressing.

Table 6.8.1 : Prefix lengths for classful addressing

Class	Prefix length	Class	Prefix length
A	/ 8	D	/ 4
B	/ 16	E	/ 4
C	/ 24		

6.8.7 Subnetting :

- The concept of subnetting in classless addressing domain is similar to that discussed for the classful addressing.
- The subnetting is used for creating a three level hierarchy in the classless addressing domain.
- An organization or an ISP have a block of addresses granted to them.
- It can divide these addresses into several subgroups and each subgroup of addresses is assigned to a **subnetwork or subnet.**
- The subnetworks may be subdivided further if the organization want it that way.

6.8.8 Designing Subnets :

Let N = Total number of addresses granted to an organization.

n = Prefix length

N_{sub} = Assigned number of addresses to each subnetwork

N_{sub} = Prefix length for each subnetwork

S = Total number of subnetworks.

- Now follow the steps given below to ensure that the subnetworks operate properly.

Steps to follow :

- The number of addresses in each subnetwork should always be equal to a power of 2. i.e. $2^0, 2^1, 2^2$....etc.
- We can use the following expression to find the prefix length of each sunetwork.

$$n_{sub} = n + \log_2 \left[\frac{N}{N_{sub}}\right] \quad ...(6.8.5)$$

- The starting address in each subnet should be divisible by the number of addresses in that subnework.
- To achieve this we need to first assign address to larger networks.

Note : These restrictions are similar to those applied when addresses to network were allocated.

6.8.9 Finding Information about Each Network :

- After designing the subnetworks, we can find the information about the subnets such as starting and last addresses, we can use the same procedure that was used to find the information about each network in the Internet.

Ex. 6.8.6 : A block of addresses granted to an ISP is given by 130.34.13.64 / 26. These addresses are to be divided into four subnetworks with equal number of hosts. Design the subnetworks and obtain all the information about each subnet.

Soln. :

Step 1 : Find total number of addresses (N) :

- From the given address we get n = 26 (prefix length).
- Hence the number of addresses in the whole network will be :

$$N = 2^{(32-n)} = 2^{(32-26)} = 2^6 = 64$$

- The first address in this block will be 130.34.13.64 / 26 whereas the last address will be 130.34.13.127 / 26.
- These values have been obtained using the procedure that we have discussed earlier.

Subnet design :

Step 2 : Find number of hosts per subnetwork :

- There are four subnetworks with equal number of guests.

∴ Number of hosts per subnetwork is given by,

$$N_1 = N_2 = N_3 = N_4 = \frac{N}{4} = \frac{64}{4} = 16 \quad \text{...Ans.}$$

- Note that the first requirement that 64 / 16 should be a power of 2 has been satisfied here.

Step 3 : Find the prefix lengths of the subnets :

- The prefix lengths of the four subnets are given by,

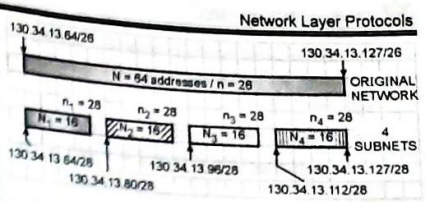
$$n_1 = n_2 = n_3 = n_4 = n + \log_2 \left[\frac{N}{N_{sub}}\right]$$

$$= 26 + \log_2 \left[\frac{64}{16}\right] = 26 + \log_2 4$$

$$\therefore n_1 = n_2 = n_3 = n_4 = 28 \quad \text{...Ans.}$$

Step 4 : Starting and ending addresses of all the subnets :

- Refer Fig. P. 6.8.6 which shows all the starting and ending addresses of the 4-subnets.



(G-1814) Fig. P. 6.8.6

- It should be noted from Fig. P. 6.8.6 that all the starting addresses should be divisible by the number of addresses in the subnet i.e. by 16.

6.8.10 Address Aggregation :

- Address aggregation is considered to be one of the advantages of CIDR architecture.
- As we know, ICANN assigns a large block of addresses to an ISP which is divided into smaller subnets and assigned to the customers by the ISPs.
- Thus many blocks of addresses are aggregated in one block and assigned to one ISP.

Ex. 6.8.7 : A router has following CIDR entries in its routing table :

Address/Mask	Next Hop
135.46.56.0/22	Interface 0
135.46.60.0/22	Interface 1
192.53.40.0/23	Router 1
Default	Router 2

For each of the following IP addresses, what does the router do if a packet with that address arrives ?

1. 135.46.63.10 2. 192.53.56.7

MU : Dec. 10, 10 Marks

SPPU : Dec. 11, 8 Marks, May 16, 5 Marks

Soln. :

CIDR – Classless Inter Domain Routing :

- IP is being heavily used for decades. However, due to the exponential growth of internet, IP is running out of addresses.
- This is a potential disaster and the internet community has begun discussion over it.
- In this section we are going to discuss one of the solutions to this problem.
- One of the solutions is CIDR (Classless Inter Domain Routing).

- The CIDR is based on the principle of allocating the remaining IP addresses in variable-sized blocks regardless of the class.
- If a site needs say 2000 addresses, then a block of 2048 addresses on the 2048 byte boundary is given to it.
- However the classless routing makes forwarding of packets more complicated.

Forwarding algorithm in the old classful system :

- The steps followed in the old classful system for forwarding packets is as follows :
  1. As soon as a packet arrives at a router, a copy of the IP address was shifted right by 28 bits to obtain a 4 bit class number.
  2. A 16-way branch then sorts packets into class A, B, C and D (if supported) with eight of the cases for class A, four of the cases for class B, two of the cases for class C and one each for D and E.
  3. The code for each class then masked off the 8–, 16–, or 24– bit network number and right aligned it in a 32 bit word.
  4. The network number was then searched in the A, B or C table.
  5. As soon as the entry was found, the outgoing line was decided and the packet was forwarded upon it.

Forwarding with CIDR :

- The simple forwarding algorithm explain earlier does not work with CIDR.
- Instead now each router table entry is extended by giving if a 32 bit mask. So now there is a single routing table for all networks (no different tables for class A, B, C, etc.) which consists of an array of triples. Each triple consists of an **IP address, subnet mask** and **outgoing line.**
- When a packet arrives at the input, the router first extracts its destination IP address. Then the routing table is scanned entry by entry to look for a match.
- It is possible that different entries with different subnet mask lengths match. In such a case the longest mask is used. For example if there is a match for a/20 mask and a/24 mask then /24 entry is used.

Solution of problem :

- Convert the IP address to bits and then AND it with the subnet mask of the interface whose address is closest to that of the IP addresses.
- The result of the ANDing will give you the network address and the interface to send the packet to.

1. IP = 135.46.63.10 :

- The interface whose address is closest to this IP is interface 1.
- This interface uses a 22 bit mask. So AND the given IP address with a 22 bit mask as follows :

IP = 135.46.63.10 = 10000111.00101110.00111111.00001010
22 bit mask = 255.255.252.0 = 11111111.11111111.11111100.00000000
IP AND Mask = 10000111.00101110.00111100.00000000
∴ IP AND Mask = 135.46.60.0

(G-1973)

- This result of ANDing matches with the network address of interface 1. Hence the router will forward this packet to interface 1.

2. IP = 192.53.56.7 :

- The interface whose address is closest to this IP is interface 2.
- This interface uses a 23 bit mask. So AND the packet IP address with a 23 bit mask as follows :

IP = 192.53.56.7 = 11000000.00110101.00111000.00000111
23 bit mask = 255.255.254.0 = 11111111.11111111.11111110.00000000
IP AND Mask = 11000000.00110101.00111000.00000000
= 192.53.56.0

(G-1974)

- This result of ANDing does not match with the network addresses of interface 0 or 1.
- Hence the packet will forwarded to the default i.e. Router 2.

6.9 Special Addresses :

- In the classful addressing, some addresses were reserved for special purpose.
- Similarly in the classless addressing as well some addresses are reserved.

6.9.1 Special Blocks :

- Some address blocks have been reserved for special purpose.

6.9.2 All Zeros Address :

- The block 0.0.0.0 / 32 contains only one address. It is called as the all zero address and has a prefix length of n = 32.
- This address has been reserved for communication when a host has to send an IPv4 packet but it does not know its own address.
- In such situations, the host sends an IPv4 packet to a DHCP server using this all zero address as the source address and a limited broadcast address (all one address) as the destination address, so as to find its own address.

6.9.3 All One Address-Limited Broadcast Address :

- The block 255.255.255.255 / 32 contains only one address.
- It is called as an all one address and has a prefix length of n = 32.
- This all one address has been reserved for limited broadcast address i.e. if a host wants to send message to all the hosts simultaneously then the sending host can use all one address as a destination address inside the IPv4 packet.
- Such a broadcasting is confined to the network only because routers do not allow the all one packet to pass through them.
- The datagram sent with the all zero address as destination will be received and processed by all the hosts on the network.

6.9.4 Loopback Address :

- A loopback address is the address which is used to test the software on a machine.
- The block 127.0.0.0 / 8 with a prefix length of 8 is used for the loopback address.
- On using this address, a packet does not leave the machine at all but it returns to the protocol software. It can be used for testing the IPv4 software.

6.9.5 Private Addresses :

- The address blocks that are not recognized globally still assigned for private use are known as private addresses.
- These addresses are neither connected to nor isolated from the Network Address Translation (NAT) techniques.
- Table 6.9.1 depict such address blocks.

Table 6.9.1 : Private addresses

Block	Number of addresses	Block	Number of addresses
10.0.0.0 / 8	16,777,216	192.168.0.0 / 16	65,536
172.16.0.0 / 12	1,047,584	169.254.0.0 / 16	65,536

6.9.6 Multicast Addresses :

- The block 224.0.0.0 / 4 with a prefix length of n = 4 has been reserved for the multicast IP communication.

6.9.7 Special Addresses in Each Block :

- The usage of some address in each block for special addresses has been recommended.
- But it has not been made mandatory. These addresses are not assigned to any host.
- One important point to be remembered is that a very small block of addresses should not be used as special addresses.

6.9.8 Network Address :

- The network address is defined as the first address (with the suffix set all to 0s) in a block.
- It is used for defining the network itself. It does not define any host in the network.
- With the same principle, the first address in a subnetwork is called as the subnetwork address.

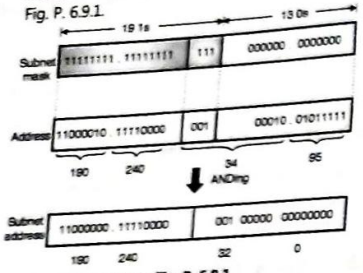
6.9.9 Direct Broadcast Address :

- We can use the last address in a block or subblock (with the suffix part set to all 1s), as a **direct broadcast address** for that block or subblock.
- A router generally uses this address for sending a packet to all the hosts connected to a specific network.
- This address is used as the destination address in the IPv4 packet and all the hosts will accept and process the datagram which has this destination address.

Ex. 6.9.1 : A router inside an organization receives the same packet with a destination address 190.240.34.95. If the subnet mask is /19 (first 19-bits are 1s and following bits are 0s). Find the subnet address.

Soln. :

- To find the subnet address, AND the destination address with the subnet mask as shown in Fig. P. 6.9.1.



(G-544) **Fig. P. 6.9.1**

- Thus the subnet address is 190.240.32.0.

Ex. 6.9.2 : An organization is granted the block 211.17.180.0/24. The administrator wants to create 32 subnets.

(a) Find the subnet mask.

(b) Find the number of addresses in each subnets.

(c) Find the first and last addresses in subnet 1.

(d) Find the first and last addresses in subnet 32.

Soln. :

Step 1 : Subnet mask :

- This is a class C network. So the default mask is given by,

255.255.255.0

- As we need 32 subnets we need 5 extra 1's. So the subnet mask will be as follows in the binary form.

Ex. 6.9.3 : An ISP is granted a block of addresses starting with 150.80.0.0/16.

The ISP wants to distribute these blocks to 2600 customers as follows :

a. The first group has 200 medium-size businesses ; each needs 128

b. The second group has 400 small businesses ; each needs 16

c. The third group has 2000 households ; each needs 4 addresses. Design the subblocks and give the slash notation for each subblock. Find out how many addresses are still available after these allocations ?

May 16, 10 Marks

Soln. :

Granted address : 150.80.0.0/16

- As n = 16 the total number of available addresses is $2^{32-n} = 2^{16} = 65536$.
- The three groups are as follows :

Group 1 :

- For this group each business needs 128 addresses.
- This means that 7 bits ($\log_2 128 = 7$) are required to define each host. The prefix length is then 32 – 7 = 25 i.e. $n_1 = 25$.
- The addresses in group-1 are

1st business : 150.80.0.0/25 to 150.80.0.127/25

2nd business : 150.80.0.128/25 to 150.80.0.255/25

.

.

.

200th business : 150.80.99.128/25 to 150.80.99.255/25

Total addresses in group-1 = 200 × 128 = 25600

Group 2 :

- For this group each business needs 16 addresses.
- Therefore 4 bits ($\log_2 16 = 4$) are required to define each host.
- The prefix length is then 32 – 4 = 28. i.e. $n_2 = 28$.
- The addresses in group-2 are

1st business : 150.80.100.0/28 to 150.80.100.15/28

2nd business : 150.80.100.16/28 to 150.80.100.31/28

.

.

400th business : 150.80.124.240/28 to 150.80.124.255/28

- Total addresses in group-2 = 400 × 16 = 6400

Group 3 :

- For this group each household needs 4 addresses.
- Therefore only 2 bits ($\log_2 4 = 2$) are required to define each host.
- The prefix length is then 32 – 2 = 30 i.e. $n_3 = 30$
- The addresses of group-3 are

1st household : 150.80.125.0/30 to 150.80.125.3/30

2000th household : 150.80.156.60/30 to 150.80.156.63/30

Total addresses in group-3 = 2000 × 4 = 8000

Number of granted address to ISP = 65, 536

Number of allocated addresses by

ISP = 25600 + 6400 + 8000 = 40,000

Number of available addresses = 65, 536 – 40,000 = 25,536.

Ex. 6.9.4 : An ISP is granted a block of addresses starting with 190.100.0.0/16 (65,536 addresses). The ISP needs to distribute these addresses to three groups of customers as follows :

1. The first group has 64 customers ; each needs 256 addresses.
2. The second group has 128 customers ; each needs 128 addresses.
3. The third group has 128 customers ; each needs 64 addresses. Design the sub blocks and find out how many addresses are still available after these allocations.

Dec. 16, 10 Marks

Soln. :

Group 1 :

- Each customer in this group needs 256 addresses i.e. suffix length is 8 ($2^8 = 256$).

∴ Prefix length = 32 – 8 = 24. The addresses are as follows :

Customer	Starting address	Ending address
1.	190.100.0.0/24	190.100.0.255/24
2.	190.100.1.0/24	190.100.1.255/24
3.	190.100.2.0/24	190.100.2.255/24
.		
.		
.		
64	190.100.63.0/24	190.100.63.255/24

Total : 64 × 256 = 16384

Group 2 :

- Each customer in this group needs 128 addresses i.e. suffix length is 7 ($2^7 = 128$).

∴ Prefix length = 32 – 7 = 25. The addresses are as follows :

Customer	Starting address	Ending address
1	190.100.64.0/25	190.100.64.127/25
2.	190.100.64.128/25	190.100.64.255/25
.		
.		
.		
128	190.100.127.128/25	190.100.127.255/25

Total : 128 × 128 = 16384

Group 3 :

- Each customer in this group needs 64 addresses i.e. suffix length is 6. ($2^6 = 64$).

∴ Prefix length = 32 – 6 = 26. The addresses are as follows :

Customer	Starting address	Ending address
1	190.100.128.0/26	190.100.128.63/26
2.	190.100.128.64/26	190.100.128.127/26
.		
.		
.		
128	190.100.159.192/26	190.100.159.255/26

Total = 128 × 64 = 8192

Number of granted addresses to the ISP = 65536

Number of allocated addresses by the ISP = 40960

Number of available addresses = 65536 – 40960 = 24576

6.10 NAT – Network Address Translation :

MU : May 19

University Questions

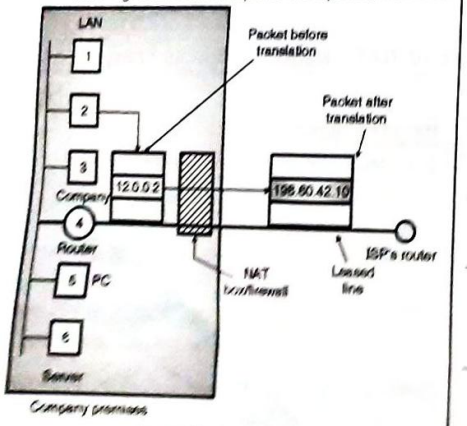
Q. 1 Short note on : Network Address Translation (NAT). **(May 19, 5 Marks)**

- The problem that existing number of IP addresses is less than the actually required ones is practically important.
- A long term solution to this problem is that the whole Internet should be migrated from IPv4 to IPv6.
- This has begun, but will take year to get complete. (That means all the computers should have IPv6 addresses instead of IPv4 addresses).

- A quick solution to this problem is NAT i.e. Network Address Translation. It is described in RFC 3022.
- The basic idea in NAT is that each company is assigned a single IP address or at the most a small number of IP addresses so as to access the Internet.
- Within the company, every computer gets a unique IP address which is used for routing the internal traffic of the office.
- But when a packet goes out of the company, and goes to ISP, the translation of IP address takes place there.
- In order to make this scheme work, three ranges of IP addresses have been declared as private.
- Companies can use these addresses internally as per their requirement.
- However no packet containing these addresses is allowed to appear on the Internet.
- The three reserved ranges are as follows :

Range1	10.0.0.0 to 10.255.255.255/8	16777216 Hosts
Range2	172.16.0.0 to 173.31.255.255/12	1048 576 Hosts
Range3	192.168.0.0 to 192.168.255.255/16	65 536 Hosts

- Generally most companies choose the addresses from the first range.
- Refer Fig. 6.10.1 which explains the operation of NAT.



(G-555) **Fig. 6.10.1 : NAT**

- It shows that within the company premises, every machine has a unique address of the form 12.a.b.c.
- But when a packet leaves the company premises, it passes through the NAT box.
- This box converts the internal IP address 12.0.0.2 in Fig. 6.10.1 to the company's true IP address 198.60.42.10.
- The NAT box is generally combined with a firewall. It is also possible to integrate the NAT box into company's router.

6.11 ICMPv4 (Internet Control Message Protocol) : MU : Dec. 16, Dec. 18, May 19

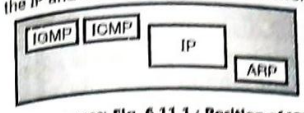
University Questions

Q. 1 What is ICMP protocol ? Explain the ICMP header format with diagram. **(Dec. 16, Dec. 18, 10 Marks)**

Q. 2 Short note on : ICMP. **(May 19, 5 Marks)**

- The IP provides unreliable and connectionless datagram delivery, and makes an efficient use of network resources.
- IP is a best-effort delivery (which does not provide any guarantee) service that takes a datagram from its original source to its final destination. However, IP has two drawbacks :
  1. It does not have any error control mechanism.
  2. It does not have any assistance mechanism.
- The Internet Control Message Protocol (ICMP) is used to overcome these drawbacks.
- It is used alongwith IP. It reports presence of errors and sends the control messages on behalf of IP.
- ICMP does not attempt to make IP a reliable protocol.
- It simply attempts to report errors and provide feedback on specific conditions.
- ICMP messages are carried as IP packets and are therefore unreliable. ICMP is a network layer protocol.
- IP also lacks a mechanism for host and management queries.
- A host sometimes wants to know if a router or another host is operating or dead.

- And sometimes a network manager needs information from another computer on the network (such as host or router).
- Fig. 6.11.1 shows the position of ICMP with respect to the IP and other protocols in the network layer.

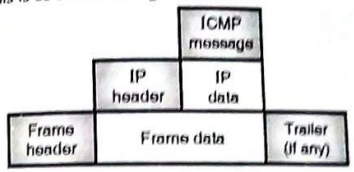


(G-2102) **Fig. 6.11.1 : Position of ICMP**

- There are two versions of ICMP protocol namely ICMPv4 and ICMPv6.
- In the following sections, we are going to discuss ICMPv4.

6.11.1 ICMP Encapsulation :

- ICMP operates in the network layer but its messages are not passed directly to the data link layer.
- Instead, the messages are first encapsulated inside IP datagrams and then sent to the lower layer.
- This is as shown in Fig. 6.11.2.



(G-2103) **Fig. 6.11.2 : ICMP encapsulation**

- The ping command uses ICMP as a probe to test whether a station is reachable.
- Ping packages an ICMP echo request message in a datagram and sends it to a selected destination.
- The user chooses the destination by specifying its IP address or name on the command line in a form such as :

ping 100.50.25.1

- When the destination receives the echo request message, it responds by sending an ICMP echo reply message.
- If a reply is not returned within a set time, ping resends the echo request several more times.
- If no reply arrives, ping indicates that the destination is unreachable.
- Another utility that uses ICMP is trace route, which provides a list of all the routers along the path to a specified destination.

6.11.2 ICMP Messages :

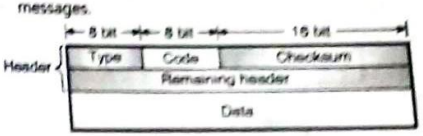
- ICMP messages are of two types :
  1. Error reporting messages
  2. Query messages.
- If a host or a router encounters a problem after processing an IP problem, then it uses the **error reporting messages** for reporting the problem.
- A host or a network manager can use the **query messages** to get some specific information from a router or another host.

6.11.3 Message Format : MU : Dec. 16, Dec. 18

University Questions

Q. 1 What is ICMP protocol ? Explain the ICMP header format with diagram. **(Dec. 16, Dec. 18, 10 Marks)**

- Fig. 6.11.3 shows the general format of ICMP messages.



(G-2105) **Fig. 6.11.3 : General format of ICMP messages**

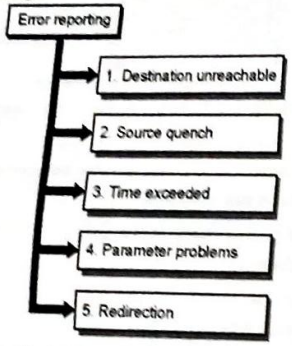
- As shown in Fig. 6.11.3, the header of an ICMP message is 8-byte long and the data section is of a variable size.
- The general header format for each ICMP message is different.
- But the first four bytes are common to all the message types.

1. **Type :**
   - This 8-bit field is used for defining the types of message.
2. **Code :**
   - This 8-bit field is used for specifying the reason for the particular message type.
   - The last common field is the **checksum** field which is 16 bit (2 byte) long. We will discuss it later in this chapter.

- The information to find the original packet that had error is included in the **data section** of the error messages.
- Whereas the **data section** in the query messages contains extra information depending on the type of query.

6.12 Error Reporting Messages (ICMPv4) :

- One of the important responsibilities of ICMP is to report the presence of an error. IP is an unreliable protocol.
- So error checking and control are not done by IP.
- So ICMP was designed to assist IP.
- But ICMP does not correct the errors. It simply reports them and leaves the error correction job to the higher level protocols.
- ICMP always sends the error reporting messages back to the original source.
- ICMP has five types of error reporting messages.
- Fig. 6.12.1 shows different types of error reporting messages.

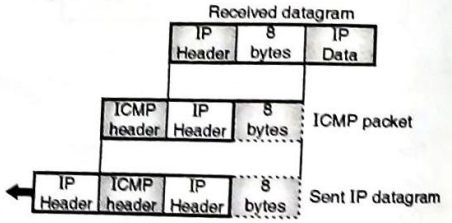


(G-2104) **Fig. 6.12.1 : Error reporting messages**

- ICMP makes use of the source IP address for sending the error message back to original source of erroneous datagram.
- Some of the important points about ICMP error messages are as follows :

1. If a datagram containing and ICMP error message is received, then no ICMP error message will be generated in response to it.
2. An ICMP error message will not be generated for a fragmented datagram that is not the first fragment.
3. Any ICMP error message will not be generated for a datagram which has a multicast address.
4. An ICMP error message will not be generated for a datagram which has a special address such as 127.0.0.0 or 0.0.0.0

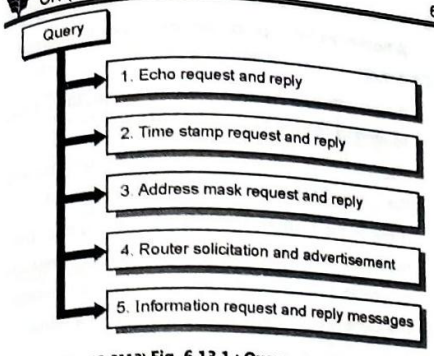
- It is important to note that the data section of every error message, contains the IP address of the original datagram in addition to the 8 bytes of data in that datagram.
- The header of the original datagram is included in the error message, to ensure that the error message will reach the original source.
- The additional 8 byte data is included because in TCP and UDP, the first 8 bytes of information contains information about the port numbers for TCP and UDP and sequence number for TCP.
- The source can use this information and convey to TCP and UDP protocols that an error has occurred.
- Then the **error packet,** formed by ICMP, is encapsulated in an IP datagram as shown in Fig. 6.12.2.



(G-2106) **Fig. 6.12.2 : Data field contents for the error message**

6.13 Query Messages (ICMPv4) :

- The ICMP can diagnose some of the network problems.
- This is in addition with the error reporting feature.
- Such a diagnosis is done through the query messages.
- The query messages is a group of five different pairs of messages as shown in Fig. 6.13.1.



(G-2112) **Fig. 6.13.1 : Query messages**

- However out of these five pairs of messages, only two pairs are being used today.
- They are :

1. Echo request and reply.
2. Timestamp request and reply.

6.14 IGMP (Internet Group Management Protocol) :

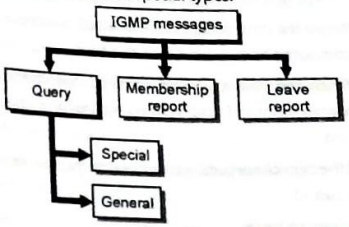
- IGMP is a necessary but not sufficient protocols used in multicasting environment. It is always used along with IP.

Group management :

- In the multicasting environment we need to use the multicast packets. So in an Internet we have to use the routers which can route multicast packets.
- A multicast routing protocol should be used to update the routing tables of these routers.
- But note that IGMP is not a multicasting routing protocol. Instead its job is to manage the group membership.
- It helps the multicast routers to create and update a list of loyal members related to each router interface.

6.14.1 Messages :

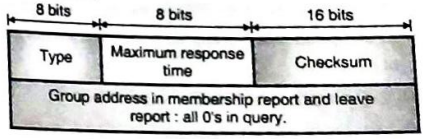
- There are two versions of IGMP called IGMPv1 and IGMPv2. In the second version there are three types of messages.
- The IGMP messages are shown in Fig. 6.14.1.
- Fig. 6.14.1 shows that there are three types namely query message, membership report and leave report.
- The query message has been divided into two types namely general and special types.



(G-586) **Fig. 6.14.1 : IGMP message types**

Message format :

- The IGMP (version-2) message format is shown in Fig. 6.14.2.



(G-587) **Fig. 6.14.2 : IGMP message format**

1. **Type :**

- It is an 8 bit field that defines the type of message as given in Table 6.14.1.
- The type and its value in hexadecimal and binary notation have also been shown in Table 6.14.1.

Table 6.14.1 : IGMP type field

Type	Value
General or special query	0x11 or 0001 0001
Membership report	0x16 or 0001 0110
Leave report	0x17 or 0001 0111

2. **Maximum response time :**

- This is the next 8-bit field which defines the amount of time allowed to answer a query.
- The value in this field shows the maximum response time in tenths of seconds.
- This value is a non zero number if the message is a query message and it is equal to zero for the other two message types.

3. **IGMP checksum :**

- The checksum is the 16-bit one's complement of the one's complement sum of the 8-byte IGMP message.

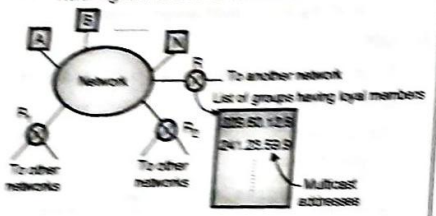
- When the checksum is computed, the checksum field should first be cleared to 0.
- When the data packet is transmitted, the checksum is computed and inserted into this field.
- When the data packet is received, the checksum is again computed and verified against the checksum field.
- If the two checksums do not match then an error has occurred.

4. **Group address :**

- This is a 32-bit field and its value depends on the type of message. For example the value of this field is zero for a general query message.
- The value in this field defines the multicast address of the group called **groupid**, in the other three types of messages.

6.14.2 Operation of IGMP :

- Refer Fig. 6.14.3 to understand the IGMP operation.

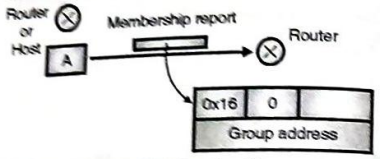


(G-588) **Fig. 6.14.3 : IGMP operation**

- IGMP operates locally. As shown in Fig. 6.14.3 multicast router R has a list of multicast addresses of the groups for which the router distributes packets.
- These packets are distributed to groups with at least one loyal member in that network.
- There is one router per group. Its duty is to distribute the multicast packets which are supposed to reach that group.
- So if there are three multicast routers (R_1, R_2 and R) connected to the network, then the lists of group identifications (ids) of all the routers are mutually exclusive i.e. they do not contain the same addresses.

6.14.3 How to Join a Group ?

- Who can join a group ? The answer is a host or a router can join a group.
- A host has a list of processes which are members of a group.
- If a processes wants to join a new group, then it has to send its request to do so to the host.
- The host adds the name of requesting process and the name of the group it wants to join to his list.
- If this is the first entry to this particular group, then the host sends a membership report message otherwise it is not necessary to send such a message.
- The same procedure is followed in case of a router.
- The process of joining a group with the format of membership report is shown in Fig. 6.14.4.



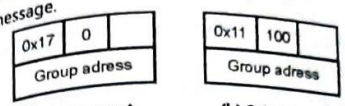
(G-589) **Fig. 6.14.4 : Membership report**

- The membership report needs to be sent twice one after the other within a quick succession, so that even if the first one is lost or damaged, the second one can be used.

6.14.4 How to Leave a Group ?

- If no process of a group is interested in a specific group then a host sends a leave report.
- Similarly if a router understands that none of the networks connected to its interfaces is interested in a particular group then it sends a leave report about that group.
- On receiving the leave report the multicast router sends a special query message and inserts the multicast address or groupid which specifies the particular group.
- The router then allows some time for any host or router to answer to this query message.
- During this time if no interest in the form of membership report is received from any one then, the router purges the group from its list.
- This is the mechanism to leave a group.

- Fig. 6.14.5(a) shows the format of the leave report while Fig. 6.14.5(b) shows the format of special query message.

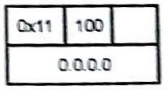


(a) Leave report **(b) Special query**

(G-2849) **Fig. 6.14.5**

6.14.5 Monitoring Membership :

- Imagine a situation in which a host is interested in a particular group but that host has been shut down or removed from the group.
- Then the multicast router will never receive any leave report.
- To handle this situation, the router periodically sends a general query message after every 125 seconds.
- The general query message does not intend to define a specific group.
- So the group address field contents should be 0.0.0.0, and the router will expect an answer for each group in its list. Even the new groups are allowed to respond.
- The maximum response time allowed for this message is 10 sec.
- The query message should be sent only by one router which is normally called as a query router.
- The format of general query message is shown in Fig. 6.14.6.



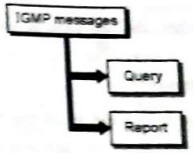
(G-2850) **Fig. 6.14.6 : General query message**

6.14.6 Query Router :

- Generally query messages create a lot of responses, which results in unnecessary traffic.
- To avoid this, the IGMP designates one router as a query router for each network.
- Only the query router is allowed to send the query message.
- All the other routers are passive. They reject responses and update their lists.

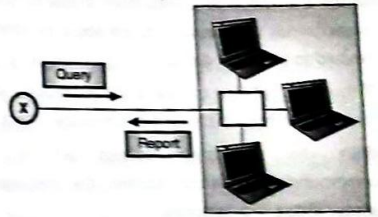
6.14.7 IGMP Messages :

- Today, for collection of information about group membership the IGMP (Internet group management protocol) is used.
- IGMP is one of the auxiliary protocol defined at the network layer which is considered as a part of Internet protocol.
- IGMP messages are encapsulated in datagram similar to ICMP messages.
- Fig. 6.14.7 shows two types of messages in IGMP version 3.



(G-2241) **Fig. 6.14.7 : IGMP Messages**

- Fig. 6.14.8 shows the operation of IGMP :



(G-2242) **Fig. 6.14.8 : Operation of IGMP**

1. **Query message :**

- A router sends query message periodically to all hosts which are attached to it to ask them for reporting their membership interest in group.
- In IGMPv3, a query message can be in any one of three forms :
  (a) a general query message,
  (b) a group specific query message,
  (c) a source and group specific message.

(a) General query message :

- In any group a general query message is sent about membership.
- With the destination address 224.0.0.1, a general query message is encapsulated in a datagram.
- All routers which are connected to the same network receive this general query message and inform them

about the message which is already sent and avoid them from resending the message.

(b) Group specific query message :

- To ask about a specific group membership, this message is sent from a router.
- If router do not receive any response about specific group and router want to make sure about a membership of that group in the network, then this group specific query message is sent. Multicast address (group identifier) is mentioned in the message.
- In a datagram, the message is encapsulated with destination address set to the corresponding multicast address.
- This message is received by all the hosts and those who are not interested they will drop this message.

(c) Source and group specific query message :

- When the message comes from a specific source, router sends this message to ask about membership related to a specific group.
- If router is not able to hear a specific group related to a specific host, then again this message is sent.
- With the destination address set to the corresponding multicast address the message is encapsulated in a datagram.
- This message is received by all the host and those who are not interested they will drop this message.

2. Report message :

- To give a response to a query message host sends report message.
- Report message contains :
  1. List of records (in which each record gives the identifier of corresponding group i.e. multicast address).
  2. Addresses of all sources in which the host is interested in receiving messages.

3. Addresses of sources from which the host do not want to receive a group message.

- In a datagram, the message is encapsulated with the multicast address 224.0.0.22 which is allocated to IGMPv3.
- In IGMPv3, if any host want to join a group, it will wait to receive a query message and then host sends a report message.
- If host want to leave the group then it will not respond to a query message.
- The group is eliminated from the router database if no hosts responds to corresponding message.

6.15 IPv6 (Next Generation IP) :

MU : Dec. 07, Dec. 08, Dec. 10

University Questions

Q. 1 Write short notes on : IPv6. **(Dec. 07, 5 Marks)**

Q. 2 List 10 important features of IPv6 protocol. **(Dec. 08, 10 Marks)**

Q. 3 Write short notes on : Features of IPv6 protocol. **(Dec. 10, 5 Marks)**

- IPv6 is the next generation Internet Protocol designed as the next step of the IP version 4.
- IPv6 was designed to enable high-performance and larger address space.
- This was achieved by overcoming many of the weaknesses of IPv4 protocol and by adding several new features.

6.15.1 Advantages of IPv6 :

MU : Dec. 07, Dec. 08, Dec. 10

University Questions

Q. 1 Write short notes on : IPv6. **(Dec. 07, 5 Marks)**

Q. 2 List 10 important features of IPv6 protocol. **(Dec. 08, 10 Marks)**

Q. 3 Write short notes on : Features of IPv6 protocol. **(Dec. 10, 5 Marks)**

1. Improved header format :

- IPv6 uses an improved header format. In its header format the options are separated from the base header.
- These options are inserted when needed, between the base header and upper layer data.
- The routing process is simplified due to this modification.
- The speed of the routing process increases and the routing time is reduced.

2. Larger address space :

- IPv6 has 128-bit address, which is 4 times wider in bits is compared to IPv4's 32-bit address space. So there is a large increase in the address space.

Address space of IPv6 = (2^{128})

3. New options :

- IPv6 has increased functionality due to the addition of entirely new options that are absent in IPv4.

4. More security :

- IPv6 includes security in the basic specification.
- It includes encryption of packets (ESP: Encapsulated Security Payload) and authentication of the sender of packets (AH : Authentication Header) for enhancing the security.

5. Possibility of extension :

- The design of IPv6 is done in such a way that there is a possibility of extension of protocol if required.

6. Support to resource allocation :

- To implement better support for real time traffic (such as video conference), IPv6 includes flow label in the specification.
- With flow label mechanism, routers can recognize to which end-to-end flow the given packet belongs to.

7. Plug and play :

- IPv6 includes plug and play in the standard specification.
- It therefore must be easier for novice users to connect their machines to the network, it will be done automatically.

8. Clearer specification and optimization :

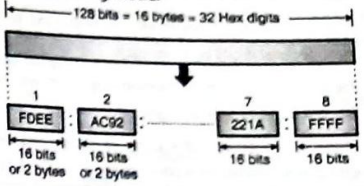
- IPv6 follows good practices of IPv4, and omits flaws/obsolete items of IPv4.

6.16 IPv6 Addressing :

- IPv6 is the next generation Internet Protocol designed as the next step of the IP version 4.
- IPv6 was designed to enable high-performance and larger address space.
- This was achieved by overcoming many of the weaknesses of IPv4 protocol and by adding several new features.
- The IPv6 was developed due to the address depletion of IPv4.
- The structure of IPv6 address is fundamentally different than that of IPv4.
- Therefore there is absolutely no possibility of address depletion taking place in future.

6.16.1 IPv6 Address :

- An IPv6 address is 128 bit long. It consists of 16 bytes as shown in Fig. 6.16.1.



(G-545) **Fig. 6.16.1 : IPv6 address**

- Thus the IPv6 address is 4 times longer than that of IPv4.

6.16.2 Notations :

- An address is stored in the computers in the binary form.
- But it is impossible for humans to handle a 128 bit binary address.
- Therefore many notations have been proposed to represent the IPv6 addresses, so that they become easier to handle for human beings.
- Some of the proposed notations are :
  1. Dotted decimal notation.
  2. Colon hexadecimal notation.
  3. Mixed representation.
  4. CIDR notation.

1. Dotted decimal notation :

- In order to maintain the compatibility with IPv4 addresses.
- We may feel tempted to use the dotted decimal notation.
- But practical observation is that this notation is convenient only for the 4 byte address of IPv4.
- It is not at all convenient for the 16 byte IPv6 addresses as it seems too long.

- Therefore this notation is very rarely used.

2. Colon hexadecimal notation :

- The 128 bit address can be made more readable and easy to handle.
- IPv6 has specified the **colon hexadecimal notation.**
- IPv6 uses a special notation called hexadecimal colon notation.
- In this, the total 128 bits are divided into 8 sections, each one is 16 bits or 2 bytes long.
- The 16 bits or 2 bytes in binary correspond to four hexadecimal digits of 4-bits each.
- Hence the 128 bits in hexadecimal form will have $8 \times 4 = 32$ hexadecimal digits.
- These are in groups of 4 digits as shown and every group is separated by a colon as shown in Fig. 6.16.2.

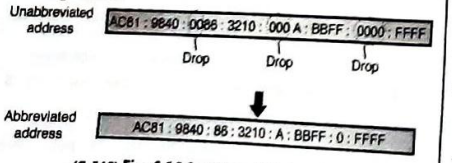
AC 81 : 9840 : 0086 : 3210 : 000A : BBFF : 0000 : FFFF

Fig. 6.16.2 : Colon hexadecimal notation

- IPv6 uses 128-bit addresses. Only about 15% of the address space is initially allocated, the remaining 85% being reserved for future use.
- These unused addresses may be used in the future for expanding the address spaces of existing address types or for totally new uses.

6.16.3 Abbreviation :

- The IPv6 address, in hexadecimal format contains 32 digits and it is very long.
- But in this address many hex digits are zero.
- We can take advantage of this to shorten the address by abbreviating it.
- A section corresponds to four digits between any two colons.
- The leading zeros in a section can be omitted to reduce the length of the address as shown in Fig. 6.16.3.

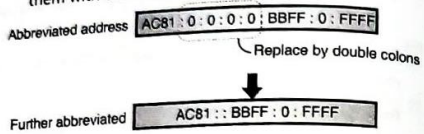


(G-546) Fig. 6.16.3 : Abbreviated address

- Note that only the leading zeros can be dropped but the trailing zeros can not be dropped. This is illustrated in Fig. 6.16.3.
- Thus due to abbreviation the length of the address has reduced to 24 hex digits from 32.

Further abbreviation :

- We can make further abbreviation if there are consecutive sections consisting of only zeros.
- This is known as **zero compression.**
- We can remove the zeros completely and replace them with double colon as shown in Fig. 6.16.4.



(G-547) Fig. 6.16.4 : Further abbreviation (Zero compression)

- This further abbreviation has reduced the address length to just 13 hex digits.
- It is important to note that abbreviation can be done only once per address.
- Also note that if there are two sets of zero sections, then only one of them can be abbreviated.

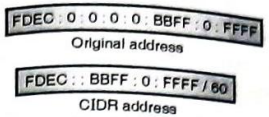
3. Mixed representation :

- Sometimes, the IPv6 address is represented using a mixed representation which combines the **colon hex** and **dotted decimal** notations.
- This notation is appropriate during the transition time during which an IPv4 address is being embedded in IPv6 address.
- In the mixed representation the rightmost 32 bits correspond to the IPv4 address.
- Hence they are represented by the dotted decimal notation.
- Whereas the leftmost 96 bits (6 sections) are represented in colon hex notation.

4. CIDR notation :

- The type of addressing used in IPv6 is **hierarchical addressing.**
- Therefore IPv6 allows classless addressing and CIDR notation.

- Fig. 6.16.5 illustrates the CIDR address with a 60 bit prefix.



(G-2132) Fig. 6.16.5 : CIDR address

- It has been discussed later on in this chapter, how we can divide an IPv6 address into a prefix and a suffix.

Ex. 6.16.1 : IPv6 uses 16-byte addresses. If a block of 1 million addresses is allocated every picosecond, how long will be the addresses last ?

Soln. :

1. Total number of address bits = $16 \times 8 = 128$
2. Number of addresses = $2^{128} = 3.4 \times 10^{38}$
3. One picosecond = 1×10^{-12} seconds
4. 1 million addresses = 1×10^{6} address

$\therefore$ 1 picosecond = 1×10^{6} addresses

$$\therefore \quad x = 3.4 \times 10^{38}$$

$$\therefore \quad x = \frac{3.4 \times 10^{38}}{1 \times 10^{6}} \times 1 \text{ picoseconds}$$

$$= 3.4 \times 10^{32} \text{ picoseconds}$$

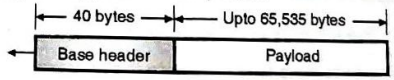
$$= 3.4 \times 10^{20} \text{ seconds} = 9.44 \times 10^{16} \text{ hours}$$

$$= 3.9352 \times 10^{15} \text{ days}$$

$$= 1.0781 \times 10^{13} \text{ years}$$

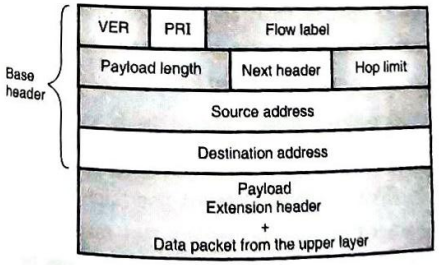
6.17 IPv6 Packet Format :

- Fig. 6.17.1(a) shows IPv6 packet.



(G-2245)Fig. 6.17.1(a) : IPv6 packet

- Fig. 6.17.1(b) shows the packet format (Base header) of IPv6.



(G-550) Fig. 6.17.1(b) : Format of an IPv6 datagram (Base header)

- Each packet can be divided into two parts viz : base header and payload.
- Base header is the mandatory part and payload is an optional one.
- The payload follows the base header.
- The payload is made up of two parts :
  1. An optional extension headers and
  2. The upper layer data.
- The base header is 40 byte long whereas the payload consisting of the extension header and upper layer data can have information worth upto 65, 535 bytes.

Base header :

- Fig. 6.17.1(b) shows the base header. It has eight fields.
- These fields are as follows :

1. Version (VER) :

- The contents of this 4 bit field defines the version of IP such as IPv4 or IPv6. If VER = 6, then the version is IPv6.

2. Priority :

- This 4 bit field contents defines the priority of the packet which is important in connection with the traffic congestion.

3. Flow label :

- It is a 24 bit (3 byte) field which is supposed to provide a special handling for a particular flow of data.

4. Payload length :

- The contents of the 16 bit or 2 byte length field are used to indicate the total length of the IP datagram excluding the base header.
- That means it gives the length of only the payload part of the datagram.

5. Next header :

- It is an 8 bit field which defines the header which follows the base header in the datagram.

6. Hop limit :

- Contents of this 8 bit (1 byte) field have the same function as TTL (time to live) in IPv4.

7. Source address :

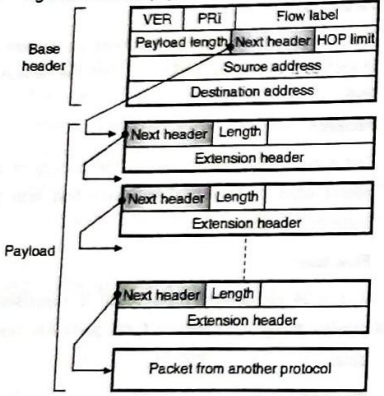
- It is a 16 byte (128 bit) Internet address which corresponds to the originator or source which has produced the datagram.

8. Destination address :

- This is a 16 byte (128 bit) internet address which corresponds to the address of the final destination of datagram.
- But this field will contain the address of the next router and not the final destination if source routing is being used.

6.17.1 Payload :

- The meaning and format of payload field in IPv6 is different as compared to payload field in IPv4.
- Fig. 6.17.2 shows payload field in IPv6.



(G-2246) **Fig. 6.17.2 : IPv6 payload**

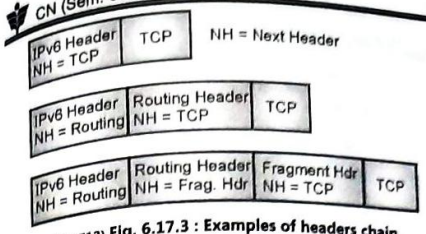
- In IPv6, the payload is combination of zero or more extension headers (options) which is followed by data from other protocols such as UDP, TCP etc.
- In IPv4, option is a part of the header, whereas in IPv6 it is designed as extension headers.
- Depends on the situation the payload can have as many extension headers as required.
- Extension header is made up of two mandatory fields : next header and the length which is followed by information which is related to the particular option.
- Value of next header field i.e. code defines which type of the next header is (e.g. source routing options, fragmentation option etc.)
- The last next header describes the protocol which carries the datagram.
- Some next header codes are listed in Table 6.17.1.

Table 6.17.1 : Next header codes

Sr. No.	Code	Next header code
1.	00	HOP by hop option
2.	02	ICMPv6
3.	06	TCP
4.	17	UDP
5.	43	Source routing option
6.	44	Fragmentation option
7.	50	Encrypted security payload
8.	51	Authentication header
9.	59	Null (no next header)
10.	60	Destination option

6.17.2 Extension Headers :

- As stated earlier the length of the base header is 40 bytes and it always remains constant.
- But in IPv6, the fixed base header can be followed by upto six extension headers.
- In IPv4 these are optional headers.
- This gives more functionality to the IP datagram.
- The IPv4 header has space for some optional fields requiring a particular processing of packets.
- These optional fields are not used often, and they can deteriorate router performance because their presence must be checked for each packet.
- IPv6 replaces these optional fields by **extension headers.**
- In IPv6, optional Internet-layer information is encoded in separate headers that may be placed between the IPv6 header and the upper-layer header in a packet (see Fig. 6.17.3).



(G-2712) **Fig. 6.17.3 : Examples of headers chain**

- There are a small number of such extension headers, each identified by a distinct Next Header value.
- An IPv6 packet may carry zero, one, or more extension headers, each identified by the Next Header field of the preceding header.
- There are seven kinds of extension header :
- Extension headers are not examined or processed by any node along a packet's delivery path, until the packet reaches the node (or each of the set of nodes, in the case of multicast) identified in the Destination Address field of the IPv6 header, except for the Hop-by-Hop Options header and the Routing header.
- Therefore, extension headers must be processed strictly in the order of their appearance in the packet; a receiver must not, for expample, scan through a packet looking for a particular kind of extension header and process that header before processing all the preceeding ones.
- Each extension header has a length equal to a multiple of 64 bits (8 bytes).
- A full implementation of IPv6 must include support for the following extension headers :
- When more than one extension header is used in the same packet, it is recommended that those headers appear in the following order :

1. IPv6 header
2. Hop-by-Hop Options header
3. Destination Options header
4. Routing header
5. Fragment header
6. Authentication header
7. Encapsulating Security Payload header
8. Destination Options header
9. Upper-layer header

1. Fragmentation :

- The fragmentation in IPv6 is conceptually same as that discussed for IPv4, but the fragmentation in IPv6 takes place at a different place than that in IPv4.
- In IPv4 the fragmentation is done by the source or router, but in IPv6 the fragmentation may be carried out only by the original source.

2. Authentication and Privacy :

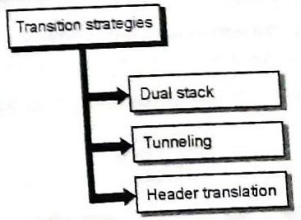
- IPv6 provides authentication and privacy using options in the extension header.

6.18 Transition from IPv4 to IPv6 :

- It is required to use a new version of the IP protocol.
- For that transition from IPv4 to IPv6 we have to define a transition day on that day each and every router or host stop using old version and should start using the new version.
- As there are huge number of systems in the Internet, transition from IPv4 to IPv6 is not practical suddenly.
- It will take some amount of time to move each and every system in the internet from IPv4 to IPv6.
- The transition from IPv4 to IPv6 should be smooth to prevent any problems in the system.

6.18.1 Transition Strategies :

- Fig. 6.18.1(a) shows the strategies for transition from IPv4 to IPv6.

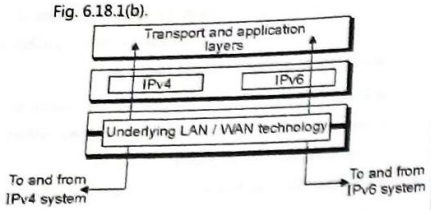


(G-2531) **Fig. 6.18.1(a) : Transition strategies**

1. Dual Stack :

- Before completely migrating to version 6 it is recommended that all hosts should have a dual stack of protocols at the time of transition.

- Simultaneously station should run IPv4 and IPv6, until the Internet uses IPv6.
- The layout of dual stack configuration is as shown in Fig. 6.18.1(b).

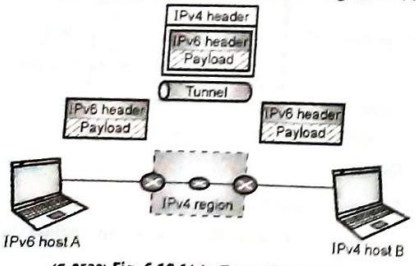


(G-2532) **Fig. 6.18.1(b) : Dual stack stategy**

- A source host send query to the DNS for deciding which version to use while sending a packet to a destination.
- A source host sends IPv4 packet if an IPv4 address is returned by the DNS, and sends IPv6 packet if DNS returns IPv6 address.

2. Tunneling :

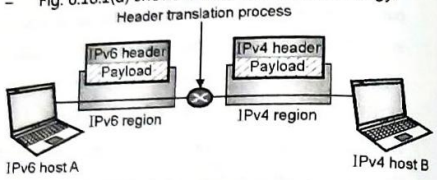
- When two computers are using IPv6 want to communicate with each other and a region through which the packet must pass uses IPv4, in such case **tunneling** strategy is used.
- The packet should have IPv4 address while passing through this region.
- When it enters in this region the IPv4 packet is encapsulated in IPv4 packet and when it exists the region it leaves its capsule.
- It looks like as if the IPv6 packet enters in a tunnel from one end and comes out from the other end.
- The protocol value is set to 41 for making it clear that IPv4 packet is holding an IPv6 packet as a data.
- The tunneling strategy is as shown in Fig. 6.18.1(c).



(G-2533) **Fig. 6.18.1(c) : Tunneling strategy**

3. Header translation :

- If some systems use IPv4 and the majority of the Internet has moved from IPv4 to IPv6, in that case header translation strategy is used where the receiver does not understand IPv6 but the sender wants to use IPv6 only.
- In this situation tunneling will not work because the packet should be in the IPv4 format which has to be understood by the receiver.
- In this strategy through header translation the format of header must be totally changed.
- The IPv6 packet header is converted into an IPv4 header.
- Fig. 6.18.1(d) shows header translation strategy.



(G-2534) **Fig. 6.18.1(d) : Header translation strategy**

6.18.2 Use of IP Addresses :

- A host may need to use both IPv4 and IPv6 addresses during the transition.
- IPv4 addresses must disappear after completion of transition.
- During the transition it is necessary that the DNS server is to be ready to map a host name to address type.
- After migrating all hosts in the world the IPv4 dictionary will disappear.

6.19 Comparison between IPv4 and IPv6 :

MU : Dec. 14

University Questions

Q. 1 Compare the network layer protocols IPv4 and IPv6 **(Dec. 14, 10 Marks)**

Table 6.19.1 : Comparison between IPv4 and IPv6

Sr. No.	IPv4	IPv6
1.	In IPv4 there are only 2^{32} possible ways to represent the address (about 4 billion possible addresses)	In IPv6 there are 2^{128} possible way (about 3.4×10^{38} possible addresses)
2.	The IPv4 address is written by dotted-decimal notation, e.g. 121.2.8.12	IPv6 is written in hexadecimal and consists of 8 groups, containing 4 hexadecimal digits or 8 groups of 16 bits each, e.g. FABC: AC77: 7834:2222:FACB: AB98: 5432:4567.
3.	The basic length of the IPv4 header comprises a minimum of 20 bytes (without option fields). The maximum total length of the IPv4 header is 60 bytes (with option fields), and it uses 13 fields to identify various control settings.	The IPv6 header is a fixed header of 40 bytes in length, and has only 8 fields. Option information is carried by the extension header, which is placed after the IPv6 header.
4.	IPv4 header has a checksum, which must be computed by each router	IPv6 has no header checksum because checksums are, for example, above the TCP/IP protocol suite, and above the Token Ring, Ethernet, etc.
5.	IPv4 contains an 8-bit field called Service Type. The Service Type field is composed of a TOS (Type of Service) field and a procedure field.	The IPv6 header contains an 8-bit field called the Traffic Class Field. This field allows the traffic source to identify the desired delivery priority of its packets
6.	The IPv4 node has only Stateful auto-configuration.	The IPv6 node has both a stateful and a stateless address autoconfiguration mechanism.
7.	Security in IPv4 networks is limited to tunneling between two networks	IPv6 has been designed to satisfy the growing and expanded need for network security.
8.	Source and destination addresses are 32 bits (4 bytes) in length.	Source and destination addresses are 128 bits (16 bytes) in length.
9.	IPsec support is optional.	IPsec support is required
10.	No identification of packet flow for QoS handling by routers is present within the IPv4 header.	Packet flow identification for QoS handling by routers is included in the IPv6 header using the Flow Label field.
11.	Address Resolution Protocol (ARP) uses broadcast ARP Request frames to resolve an IPv4 address to a link layer address.	ARP Request frames are replaced with multicast Neighbour Solicitation messages.
12.	Must be configured either manually or through DHCP.	Does not require manual configuration or DHCP.
13.	ICMP Router Discovery is used to determine the IPv4 address of the best default gateway and is optional	ICMP Router Discovery is replaced with ICMPv6 Router Solicitation and Router Advertisement messages and is required.
14.	Header includes options	All optional data is moved to IPv6 extension headers.

Review Questions

Q. 1 Name different protocols in the network layer.

Q. 2 Explain the purpose of ARP.

Q. 3 Why is ARP request broadcast but ARP reply unicast ?

Q. 4 Write a note on IP.

Q. 5 Explain fragmentation in IP.

Q. 6 What is the name of a packet in IP ?

Q. 7 Explain the IP header.

Q. 8 Compare IPv4 and IPv6.

Q. 9 State limitations of IPv4.

Q. 10 Write a note on ICMP.

Q. 11 Name and describe three types of IPv6 addresses.

Q. 12 What is fragmentation ? Explain how is it supported in IPv4 and IPv6.

Q. 13 Explain the addressing scheme in IPv4 and IPv6. When IPv6 protocol is introduced, does the ARP protocol have to be changed ? Explain.

Q. 14 Given an IP address, how will you extract its net id and host id.

Q. 15 What is subnetting in IP network, explain with suitable examples.

Q. 16 Why is an ARP Query sent within a broadcast frame ? Why is an ARP response sent within a frame with a specific destination LAN address ?

Q. 17 A network on the internet has a subnet mask of 255.255.240.0. What is the maximum number of hosts it can handle ?

Q. 18 An IP datagram using the strict source routing option has to be fragmented. Do you think the option is copied into each fragment, or is it sufficient to just put it is the first fragment ? Explain your answer.

□□□

Module 5

Chapter 7

Transport Layer

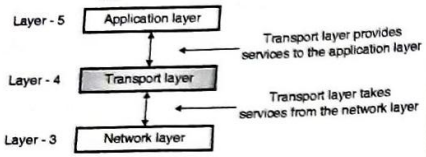
Syllabus

The Transport Service: Transport service primitives, Berkeley Sockets, Connection management (Handshake), UDP, TCP, TCP state transition, TCP timers, TCP Flow control (sliding Window), **TCP Congestion Control :** Slow Start.

Chapter Contents

7.1 Introduction :

- The transport layer is the core of the Internet model.
- The application layer programs interact with each other using the services of the transport layer.
- Transport layer provides services to the application layer and takes services from the network layer.
- Fig. 7.1.1 shows the position of the transport layer in the 5-layer internet model.

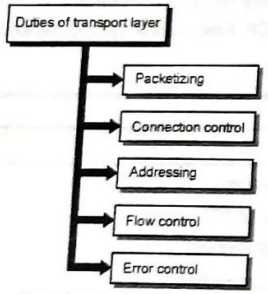


(G-592) **Fig. 7.1.1 : Position of transport layer**

- The transport layer is fourth layer in this model. It connects the lower three layers to upper three layers of an OSI layer.

7.2 Transport Layer Duties :

- Transport layer is meant for the process to process delivery and it is achieved by performing a number of functions.
- Fig. 7.2.1 lists the functions of a transport layer.



(G-1407) **Fig. 7.2.1 : Duties of transport layer**

1. Packetizing :

- The transport layer creates packets with the help of encapsulation on the messages received from the application layer. Packetizing is a process of dividing a long message into smaller ones.
- These packets are then encapsulated into the data field of the transport layer packet.
- The headers containing source and destination address are then added.
- The length of the message which is to be divided can vary from several lines (e-mail) to several pages.
- But the size of the message can become a problem.
- The message size can be larger than the maximum size that can be handled by the lower layer protocols.
- Hence the messages must be divided into smaller sections. Each small section is then encapsulated into a separate packet.
- Then a header is added to each packet to allow the transport layer to perform its other functions.

2. Connection control :

- Transport layer protocols are divided into two categories :

1. Connection oriented. 2. Connectionless.

Connection oriented delivery :

- A connection oriented transport layer protocol establishes a connection i.e. virtual path between sender and receiver.
- This is a virtual connection. The packet may travel out of order.
- The packets are numbered consecutively and communication is bi directional.

Connectionless delivery :

- A connectionless transport protocol will treat each packet independently.
- There is no connection between them. Each packet can take its own different route.

3. Addressing :

- The client needs the address of the remote computer it wants to communicate with.
- Such a remote computer has a unique address so that it can be distinguished from all the other computers.

4. Flow and error control :

- For high reliability the flow control and error control should be incorporated.
- **Flow control :** We know that data link layer can provide the flow control. Similarly transport layer also can provide flow control.
- But this flow control is performed end to end and not across a single link.

- **Error control :** The transport layer can provide error control as well.
- But error control at transport layer is performed end to end and not across a single link.
- Error correction is generally achieved by retransmission of the packets discarded due to errors.

Congestion control and QoS :

- The congestion can take place in the data link, network or transport layer.
- But the effect of congestion is generally evident in the transport layer.
- Quality of Service (QoS) can be implemented in other layers but its actual effect is felt in the transport layer.
- The transport layer enhances the QoS provided by the network layer.

7.3 Transport Layer Services :

- In this section we are going to discuss the services provided by the transport layer.

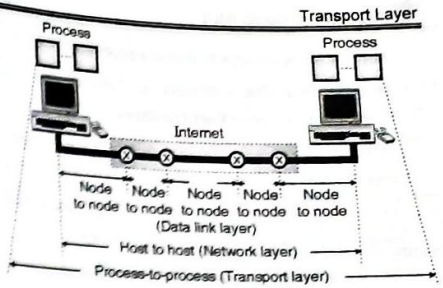
7.3.1 Process-to-Process Communication :

MU : Dec. 05, May 07

University Questions

Q. 1 Discuss the services offered by transport layer. **(Dec. 05, May 07, 8 Marks)**

- The data link layer performs a node to node delivery.
- The network layer carries out the datagram delivery between two hosts (host to host delivery).
- But the real communication takes place between two processes or application programs for which we need the **process-to-process delivery.**
- The transport layer takes care of the **process-to-process delivery.**
- In this a packet from one process is delivered to the other process.
- The relationship between the communicating processes is the client-server relationship.
- Fig. 7.3.1 demonstrates the three processes.



(G-594) **Fig. 7.3.1 : Types of data deliveries**

- There is a difference between host-to-host communication and process to process communication that we need to understand clearly.
- The host to host (computer to computer) communication is handled by the network layer.
- But this communication only ensures that the message is delivered to the destination computer. But this is not enough.
- It is necessary to handover this message to the correct process. The transport layer will take care of this.

7.3.2 Addressing : Port Number :

MU : Dec. 05, May 07

University Questions

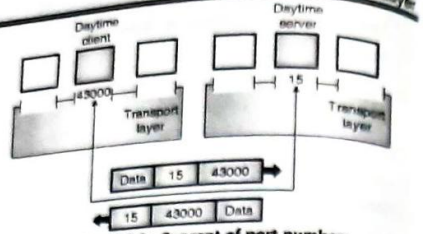
Q. 1 Discuss the services offered by transport layer. **(Dec. 05, May 07, 8 Marks)**

- There are several ways of achieving the process-to-process communication, but the most common method is using the client-server paradigm.
- **Client** is defined as the process on the local host.
- It needs services from another process called **server** which is on the other (remote) host.
- Both client and server have the same name. Some of the important terms related to the client-server paradigm are :

1. Local host
2. Remote host
3. Local process
4. Remote process

- We can use the IP addresses to define the local host and remote host.

- But this is not enough to define a process.
- In order to define a process, we have to use one more identifier called **Port Numbers.**
- In TCP/protocol suite, the port numbers are integers and they are numbered between 0 and 65,535.
- At the data link layer we need a MAC address, at the network layer we need to use an IP address.
- A datagram uses the destination IP address to deliver the datagram and uses the source IP address for the destination's reply.
- At the transport layer a transport layer address called a **port number** is required to be used to choose among multiple processes running on the destination host.
- The destination port number is required to make the packet delivery and the source port number is needed to return back the reply.
- In the Internet model, the port numbers are 16 bit integers.
- Hence the number of possible port numbers will be 2^{16} = 65,535 and the port numbers range from 0 to 65,535.
- The client program identifies itself with a port number which is chosen randomly.
- This number is called as **ephemeral port number.** Ephemeral means short lived. It is used because life of a client is generally short.
- The server process should also identify itself with a port number but this port number can not be chosen randomly.
- The Internet uses universal port numbers for servers and these numbers are called as **well known port numbers.**
- Every client process knows the well known port numbers of the pre identified server process.
- For example, a Day time client process can use an ephemeral (temporary) port number 43000 for identifying itself, the Day time server process must use the well known (permanent) port number 15.
- This is illustrated in Fig. 7.3.2.



(G-595) **Fig. 7.3.2 : Concept of port numbers**

What is difference between IP addresses and Port numbers ?

- The IP addresses and port numbers have altogether different roles in selecting the final destination of data.
- The destination IP address is used for defining a particular host among the millions of hosts in the world.
- After a particular host is selected, the port number is used for identifying one of the processes on this selected host.

IANA Ranges :

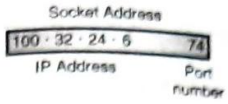
- The port numbers are divided into three ranges by IANA (International Assigned Number Authority).
- The ranges are as follows :
  1. Well known ports
  2. Registered ports
  3. Dynamic or private ports.

1. **Well known ports :** The ports from 0 to 1023 are known as well known ports. They are assigned as well as controlled by IANA.
2. **Registered ports :** The ports from 1024 to 49,151 are neither controlled nor assigned by IANA. We can only register them with IANA to avoid duplication.
3. **Dynamic or private ports :** The ports from 49,152 to 63,535 are known as dynamic ports and they are neither controlled nor registered.
   - They can be used by any process. Dynamic ports are also known as private ports and dynamic port are called as ephemeral ports.

Socket address :

- Process to process delivery (transport layer communication) has to use two addresses, one is IP address and the other is port number at each end to make a connection.

- Hence a process to process delivery uses the combination of these two.
- The combination of IP address and port number is as shown in Fig. 7.3.3 and it is known as the socket address.



(G-1548) **Fig. 7.3.3 : Socket address**

- The client socket address defines the client process uniquely whereas the server socket address defines the server process uniquely.
- A transport layer protocol requires the client socket address as well as the server socket address.
- These two addresses contain four pieces.
- These four pieces go into the IP header and the transport layer protocol header.
- The IP header contains the IP addresses while the UDP and TCP headers contain the port numbers.
- If we want to use the transport layer services in the Internet, then we have to use a pair of socket addresses namely the clients socket address and the server's socket address.

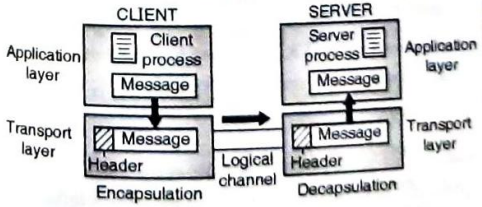
7.3.3 Encapsulation and Decapsulation :

MU : Dec. 05, May 07

University Questions

Q. 1 Discuss the services offered by transport layer. **(Dec. 05, May 07, 8 Marks)**

- The transport layer carries out the **Encapsulation** of the message at the sending end and then **Decapsulation** at the receiving end when two computers communicate.
- This process has been illustrated in Fig. 7.3.4.



(G-2012) **Fig. 7.3.4 : Encapsulation and decapsulation**

Encapsulation :

- At the sending end the process that has a message to send, will pass it to the transport layer alongwith a pair of socket addresses and some additional information.
- The transport layer adds its own header to this data. This packet at the transport layer in the Internet is known by different names such as **user datagram, segment or packet.**

Decapsulation :

- When the segment or datagram arrives at the receiving end, the header is isolated and destroyed, and the message is delivered to the process running at the application layer as shown in Fig. 7.3.4.
- The socket address of the sender process is then handed over to the destination process.

7.3.4 Multiplexing and Demultiplexing :

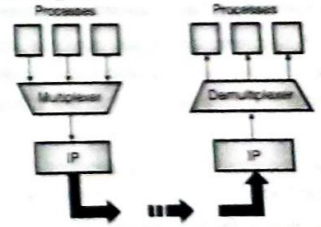
MU : Dec. 03, Dec. 05, May 07

University Questions

Q. 1 How are the port numbers used by TCP/UDP in demultiplexing incoming segments ? **(Dec. 03, 10 Marks)**

Q. 2 Discuss the services offered by transport layer. **(Dec. 05, May 07, 8 Marks)**

- The addressing mechanism allows multiplexing and demultiplexing taking place at the transport layer as shown in Fig. 7.3.5.



(G-597) **Fig. 7.3.5 : Multiplexing and demultiplexing**

Multiplexing :

- At the sending end, there are several processes that are interested in sending packets.
- But there is only one transport layer protocol (UDP or TCP).
- Thus it is a many processes-one transport layer protocol situation.

- Such a many-to-one relationship requires multiplexing.
- The protocol first accepts messages from different processes.
- These messages are separated from each other by their port numbers. Each process has a unique port number assigned to it.
- Then the transport layer adds header and passes the packet to the network layer as shown in Fig. 7.3.5.

Demultiplexing :

- At the receiving end, the relationship is one as to many. So we need a demultiplexer.
- First the transport layer receives datagrams from the network layer.
- The transport layer then checks for errors and drops the header to obtain the messages and delivers them to appropriate process based on the port number.

7.3.5 Flow Control :

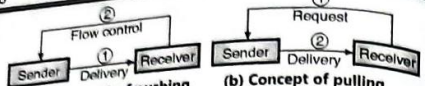
- If the packets produced by the sender are at a rate X and the receiver is receiving them at a rate Y, then for
  X = Y, there will be a perfect balance observed in the system.
- But if X is higher than Y (source is producing packets at a rate which is higher than the rate at which the receiver is accepting them), then the receiver can be overwhelmed and has to **discard** some packets.
- And if X is less than Y (i.e. source is producing packets at slower rate than the rate of acceptance at the receiver) then system becomes **less efficient.**
- Flow control is related to the situation in which X > Y because it is very important to prevent data loss (due to discarding of packets) at the receiver site.

Pushing and pulling for flow control :

- There are two different ways of delivering the packets produced by the sender to the receiver. They are pushing or pulling.

1. Pushing :

- If the sender is sending the packets soon as they are produced, without receiving any prior request from the receiver then this type of deliver is called as **pushing.** Fig. 7.3.6(a) illustrates this concept.



(a) Concept of pushing (b) Concept of pulling

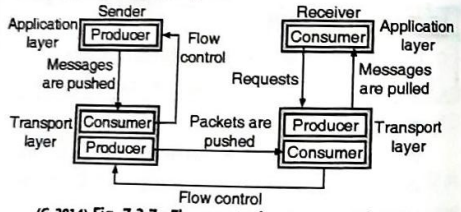
(G-2013) Fig. 7.3.6

2. Pulling :

- If the sender sends the produced packets only when they are requested by the receiver then the delivery is called as **pulling.** Fig. 7.3.6(b) illustrates the principle of pulling.
- In case of **pushing** type delivery, if the packets are being sent at a higher rate than that of receiving, then the receiver will be **overwhelmed,** and some received packets will have to be discarded.
- In order to avoid discarding of packets, the **flow control** will have to be exercised.
- For this the receiver has to warn the sender to stop the delivery when it is overwhelmed and it has to inform the sender again to start delivery when it (receiver) is ready, to receive the packets.
- In case of **pulling type delivery,** the receiver is actually pulling the packets from the sender.
- It requests for the packets when it is ready. Therefore the flow control is not required in this case.

7.3.6 Flow Control at Transport Layer :

- The concept of flow control at transport layer has been illustrated in Fig. 7.3.7.



(G-2014) Fig. 7.3.7 : Flow control at transport layer

- It shows the communication taking place between a sender and a receiver.
- As shown in Fig. 7.3.7, there are four entities involved in this communication. They are as follows :
  1. Sender process.
  2. Sender transport layer.
  3. Receiver process.
  4. Receiver transport layer.

- We will discuss the flow control by considering the sending and receiving ends separately.

Sending end :

- The first entity on the sending end is the **sender process,** at the application layer. It works only as a **producer** which produces chunk of messages and pushes them to the transport layer on the sending end, as shown in Fig. 7.3.7.
- The second entity on the sending end is the **sender transport layer.** It has two different roles to play.
- First it acts as a **customer** and consumes all the messages produced and pushed by the producer.
- Then it encapsulates those messages into packets and pushes them to the receiver transport layer as shown in Fig. 7.3.7. Here it acts as a **producer.**

Receiving end :

- The first entity on the receiving end is the **receiver transport layer.** It also has two different roles to play.
- It acts as a **consumer** for the packets pushed by the senders transport layer and it also acts as the **producer.**
- It has decapsulate the messages and deliver them to the application layer as shown in Fig. 7.3.7.
- However the delivery of decapsulated messages to the application layer is a **pulling type delivery.**
- That means the transport layer waits till the application layer process requests for the decapsulated messages.

Flow control :

- As shown in Fig. 7.3.7, the flow control is needed for atleast two cases.
- First is from transport layer of sender to the application layer of sender.
- And secondly form the transport layer of receiver to the transport layer of sender.

Buffers :

- It is possible to implement the flow control in many different ways.
- One of the ways of implementation is to use two **buffers** one each at the sending and receiving transport layers.
- A **buffer** is nothing but a set of memory locations which can temporarily hold (store) packets.
- It is possible to exercise flow control communication by sending signals from the consumer to producer.
- The **flow control** at the **sending end** takes place as follows :
- As soon as the buffer at the transport layer becomes full it sends the stop message to its application layer in order to stop the chunk of messages that are being pushed into the buffer.
- The second flow control takes place at the receiver transport layer as follows :
- As soon as the buffer at receiver transport layer becomes full, it will inform the sender transport layer to stop pushing the packets.
- Whenever the buffer becomes partially empty, it again informs the sender transport layer to start sending the packets again.

7.3.7 Error Control :

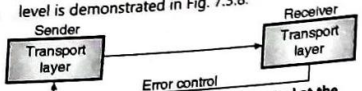
Need of error control :

- In the Internet, the network layer protocol IP has the responsibility to carry the packets from the transport layer at the sending end to the transport layer at the receiving end.
- But IP is unreliable. Therefore transport layer should be made reliable, in order to ensure reliability at the application layer.
- We can make the transport layer reliable by adding the **error control service** to the transport layer.

Duties of error control mechanism :

- Following are the important responsibilities of the error control mechanism introduced at the transport layer :
  1. To find and discard the corrupted packets.
  2. To keep the track of lost and discarded packets and to resend them.
  3. Identify the duplicate packets and discard them.
  4. To buffer out of order packets until the missing packets arrive.
- In the error control process, only the sending and receiving transport layers are involved.

- That means it is assumed that the chunk of messages exchanged between the application layers and transport layers are error free.
- The concept of error control at the transport layer level is demonstrated in Fig. 7.3.8.



(G-2015) **Fig. 7.3.8 : Concept of error control at the transport layer**

- The receiving transport layer manages the error control by communicating with the sending transport layer about the problem.

Sequence numbers :

- In order to exercise the error control at the transport layer following two requirements should be satisfied :
  1. The sending transport layer should know about the packet which is to be resent.
  2. The receiving transport layer should know about the packets which are duplicate or the ones that have arrived out of order.
- The requirements can be satisfied only if each packet has a unique **sequence number.**
- If a packet is either corrupted or lost the receiving transport layer will somehow inform the sending transport layer about the sequence number of those packets and request it to resend those packets.
- Due to the unique sequence number assigned to each packet it is possible for the receiving transport layer to identify the duplicate packets received.
- The out of order packets can also be recognized by observing gaps in the sequence numbers of the received packets.
- Packet numbers are given sequentially. But the length of the sequence number cannot be too long because the sequence number is to be included in the header of the packets.
- If the header of a packet allows "m" bits per sequence number, then the range of sequence number will be from 0 to 2^m–1.
- For example if m = 3 then the range of sequence numbers will be from 0 to 7.
- Thus sequence numbers are modulo 2^m.

Acknowledgement :

- The receiver side can send an acknowledgement (ACK) signal corresponding to each packet or each group of packets which arrived safe and sound.
- The question is what happens if a received packet is corrupted ? The answer is that the receiver simply discards the corrupted packet and does not send any ACK signal for it.
- The sender can detect a lost packet with the help of a timer. A timer is started at the sending end as soon as a packet is sent.
- If the ACK does not arrive before the expiry of the timer, then the sender treats the packet to be either lost or corrupted and resends it.
- The receiver silently discards the duplicate packets. It will either discard the out of order packets or stored until the missing packet is received.
- Note that every discarded packet is treated as a lost packet by the sender.

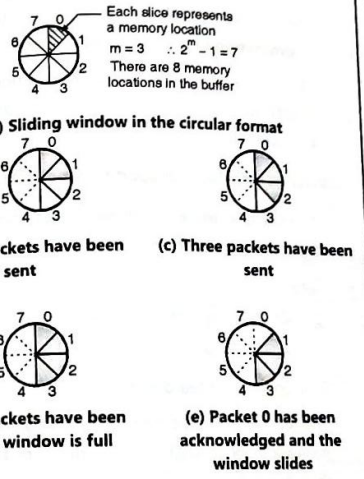
7.3.8 Combination of Flow and Error Control :

- Till now we have discussed the following important concepts :
  1. We need to use buffers at the sending and receiving ends for exercising the flow control.
  2. Also we have to use the sequence numbers and acknowledgements for exercising the error control.
- We can combine these two concepts together by using two numbered buffers one at the sender and the other at the receiver, in order to exercise a combination of flow and error control.
- At the sending end, when a packet, is prepared to be sent, the number of the next free location (x) in the buffer is used as the sequence number of that packet.
- As soon as the packet is sent, its copy is stored at location (x) in the sending end buffer and the sender waits for the acknowledgement from the receiver.
- On reception of the acknowledgement of the sent packet, the copy of that packet is purged to make the memory location (x) free again.

- At the receiver, when a packet having a sequence number "y" arrives, it is stored at the memory location "y" in the receiver buffer until the receiver application layer is ready to receive it.
- The receiver will send the ACK message back to sender to inform it that packet "y" has arrived.

Sliding window :

- As the sequence numbers are modulo 2^m, we can use a circle as shown in Fig. 7.3.9 to represent the sequence number from 0 to 2^m–1.

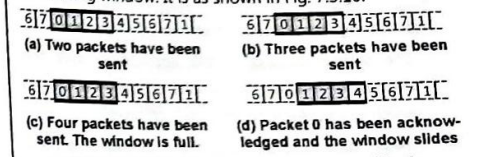


(G-2017) **Fig. 7.3.9**

- We can represent the buffer as a set of slices, called as the **sliding window** which will occupy a part of the circle at any time.
- In Fig. 7.3.9, we have assumed that m = 3. Therefore $2^m - 1 = 7$ and the sequence numbers are from 0 to 7.
- Hence the number of memory locations in a buffer will also be 8 i.e. 0 to 7.
- The sliding windows will correspond to the sender as well as receiver.
- On the sending side, when a packet is sent we will mark the corresponding slice.
- Therefore when marking of all the slices is done, it means the **sending buffer is full,** and it cannot accept any further messages from the application layer as shown in Fig. 7.3.9(d).
- When the acknowledgement for segment "0" arrives at the sending end, the corresponding segment (segment 0) is unmarked and window slides ahead by one slice as shown in Fig. 7.3.9(e).
- The size of the **sending window** is 4.
- Note that the sliding window is just an abstraction. In actual practice, computer variables are used to hold the sequence number of the next packet to be sent and the last packet sent.

Sliding window in the linear format :

- This is another way to diagrammatically represent a sliding window. It is as shown in Fig. 7.3.10.



(G-2075)**Fig. 7.3.10 : Sliding windows presented in the linear format**

- The principle of this type of sliding window is same as that of the circular representation.
- The linear format is the most preferred format. It needs less space on paper.
- Fig. 7.3.10(a), (b), (c) and (d) are the sliding windows presented in the linear format corresponding to Figs. 7.3.9(b), (c), (d) and (e) respectively in the circular presentation.

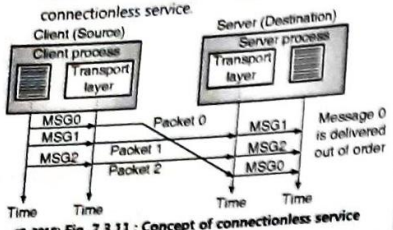
7.3.9 Connectionless and Connection Oriented Services (CLTS & COTS) :

- A transport layer protocol is capable of providing two types of services :
  1. Connectionless services.
  2. Connection oriented services.
- The meaning of the words connectionless and connection oriented is different at the transport layer than that at the network layer.
- A connectionless service at the network layer means different datagrams of the same message following different paths.
- However at the transport layer, the meaning of connectionless service is independency between different packets.

- On the other hand a connection oriented service means the packets are interdependent.

Connectionless Transport Service (CLTS) :

- Refer Fig. 7.3.11 to understand the concept of connectionless service.



(G-2018) Fig. 7.3.11 : Concept of connectionless service

- The source process at the application layer first divides its message in chunks of data the size of which is acceptable to the transport layer.
- These data chunks are then delivered to the transport layer one by one.
- These chunks are treated as independent units by the transport layer.
- Every data chunk arriving from the application layer is encapsulated in a packet by the transport layer and sent to the destination transport layer as shown in Fig. 7.3.11.

Out of Order Delivery :

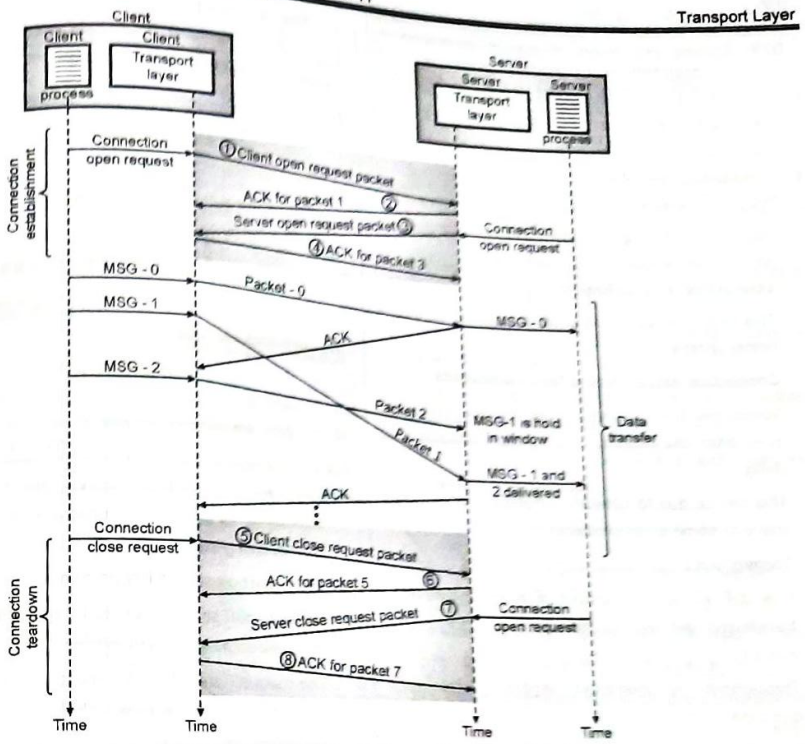
- In Fig. 7.3.11 we have considered three chunks of independent messages 0, 1 and 2.
- As the corresponding packets also are independent of each other and as they are free to follow their own path, these packets can arrive out of order at the destination as shown in Fig. 7.3.11.
- Naturally they are delivered to server process in an out of order manner.
- As seen in Fig. 7.3.11, at the sending end (client) the three chunks of messages 0, 1 and 2 are delivered to the transport layer in the order 0, 1, 2.
- But packet 0 travels a longer path and undergoes an extra delay.
- Therefore the packets are not delivered in order at the destination (server) transport layer.
- Therefore the message chunks delivered to the server process will also be out of order (1, 2, 0).
- If these chunks are of the same message then due to their out of order delivery the server will receive a strange message.

One packet is lost :

- The UDP packets are not numbered. So if one of the packets is lost, then the receiving transport layer will not have any idea about the lost packet.
- It will simply deliver the received chunks of messages to the server process.
- The above problems arise due to **lack of coordination** between the two transport layers.
- Due to this lack of co-ordination it is not possible to implement flow control, error control or congestion control in the connectionless service.

Connection Oriented Transport Service (COTS) :

- As we know, there are three stages involved in the connection oriented service.
- They are :
  1. Connection establishment.
  2. Exchange of data.
  3. Connection teardown.
- The connection oriented service is present at the network layer as well, but it is different from that at the transport layer.
- At the network layer, the meaning of connection oriented service involves the co-ordination between the hosts on either sides and all the routers between them.
- But at the transport layer, the meaning of connection oriented service is the end to end service that involves only the two hosts.
- Refer Fig. 7.3.12 to understand the concept of connection oriented service at the transport layer.
- In Fig. 7.3.12, all the three stages namely connection establishment, data exchange and connection teardown have been shown.
- It is important to note that it is possible to implement the flow control, error control and congestion control in the connection oriented service.



(G-2076) Fig. 7.3.12 : Concept of connection oriented service

Comparison of Connection Oriented and Connectionless Services :

Sr. No.	Parameter	Connection oriented	Connectionless
1.	Reservation of resources	Necessary	Not necessary
2.	Utilization of resources	Less	Good
3.	State information	Lot of information required	Not much information is required to be stored
4.	Guarantee of service	Guaranteed	No guarantee
5.	Connection	Connection needs to be established	Connection need not be established
6.	Delays	More	Less
7.	Overheads	Less	More
8.	Packets travel	Sequentially	Randomly
9.	Congestion due to overloading	Not possible	Very much possible

7.3.10 Quality of Service (QoS) :

MU : Dec. 04, Dec. 06, May 09, May 15, May 16, Dec. 16

University Questions

Q. 1 Briefly explain the primary parameters that are the requirements to provide Quality of Service in networks. **(Dec. 04, 10 Marks)**

Q. 2 Write short notes on : QoS in transport layer. **(Dec. 06, May 09, 3 Marks)**

Q. 3 Discuss the quality of service parameters in computer network. **(May 15, May 16, Dec. 16, 10 Marks)**

- As mentioned earlier, the QoS parameters are as follows :

1. Connection establishment delay :

- The time difference between the instant at which a request for transport connection is made and the instant at which it is confirmed is called as **connection establishment delay.**
- This delay should be as short as possible to ensure better service.

2. Connection establishment failure probability :

- Sometimes the connection may not get established even after the maximum connection establishment delay.
- This can be due to network congestion, lack of table space or some other problems.

3. Throughput :

- It is defined as the number of bytes of user data transferred per second, measured over some time interval.
- Throughput is measured separately for each direction.

4. Transit delay :

- It is the time duration between a message being sent by the transport user from the source machine and its being received by the transport user at the destination machine.

5. Residual error ratio :

- It measures the number of lost or garbled messages as a percentage of the total messages sent.
- Ideally the value of this ratio should be zero and practically it should be as small as possible.

6. Protection :

- This parameter provides a way to protect the transmitted data against reading or modifying it by some unauthorised parties.

7. Priority :

- Using this parameter the user can show that some of its connections are more important (have higher priority) than the other ones.
- This is important when congestions take place. Because the higher priority connections should get service before the low priority connections.

8. Resilience :

- Due to internal problem or congestion the transport layer spontaneously terminates a connection.
- The resilience parameter gives the probability of such a termination.

7.4 Transport Service Primitives :

MU : Dec. 15, Dec. 17

University Questions

Q. 1 What are transport service primitives ? Discuss in brief. **(Dec. 15, 10 Marks)**

Q. 2 What are transport service primitives ? Explain. **(Dec. 17, 10 Marks)**

- The transport service primitives allow the transport user such as application programs to access the transport service.
- Each transport service has its own access primitives.
- The transport service is similar to network service but there are some important differences.
- The main difference is that the connection-oriented transport service is reliable.
- The second difference between the network service and transport service is whom the services are intended for.
- The transport primitives are seen by many programs and programmers. Hence the transport service is convenient and easy to use.
- We can get the idea about the transport services by referring to Table 7.4.1 which lists the five primitives.

Table 7.4.1 : Primitives for a simple transport service

Sr. No.	Primitive	TPDU sent	Meaning
1.	LISTEN	None	Block until some process tries to connect
2.	CONNECT	Connection request	Actively attempt to establish a connection
3.	SEND	Data	Send data

Sr. No.	Primitive	TPDU sent	Meaning
4.	RECEIVE	None	Block until a data TPDU arrives
5.	DISCONNECT	Disconnection request	Release the connection

- The transport interface allows the application programs to establish, use and release connections.
- Let us see how these primitives are used in actual applications.

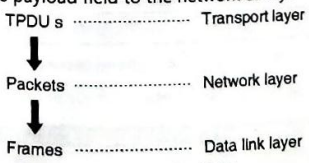
1. The server executes a LISTEN primitive. This will make a system call to block the server until a client turns up.
2. When a client wants to talk to the server it executes the CONNECT primitive.
3. In response the transport entity blocks the caller and sends a packet to the server. The transport layer message is encapsulated in the payload of this packet for the server's transport entity.

TPDU :

- The message sent from transport entity to transport entity is called as transport protocol data unit or TPDU.

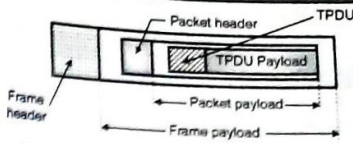
Nesting of TPDUs, Packets and Frames :

- The TPDUs which are exchanged by the transport layer are contained in the packets that are exchanged by the network layer.
- These packets are in turn contained in the frames which are exchanged by the data link layer.
- When a frame arrives, the data link layer processes the frame header and passes the contents of the frame payload field to the network entity.



(G-599(a)) **Fig. 7.4.1**

- The network entity processes the packet header and passes the contents of the packet payload to the transport entity.
- This is called as "Nesting" and it is illustrated in Fig. 7.4.2.



(G-600) **Fig. 7.4.2 : Nesting of TPDUs, packets and frames**

Connect primitive :

- If a client gives the CONNECT call, then a connection request TPDU is sent to the server.
- When this TPDU arrives, the transport entity checks if the server is blocked on a LISTEN.
- It then unblocks the server and sends a connection accepted TPDU back to the client.
- On arrival of this TPDU, the client is unblocked and connection is established.

SEND and RECEIVE Primitives :

- The SEND and RECEIVE primitives can be used for exchange of data.
- The data exchange at the network layer is more complicated than that at the transport layer.
- In transport layer data exchange, every data packet is eventually acknowledged. The packets carrying control TPDUs are also acknowledged.
- All these acknowledgements are managed by the transport entities using the network layer protocols.
- The transport entities have to take care of issues like timers and re-transmission.
- The transport layer connection acts as a reliable bit pipe through which the bits sent by a sender come out from the other side of pipe.

Connection release :

- A connection should be released when it is no longer needed.
- This is essential in order to free up the table space within the two transport entities.
- Disconnection can be of two types :
  1. Asymmetric
  2. Symmetric

7.5 Sockets :

MU : Dec. 03, Dec. 06, Dec. 07, May 08, May 09

University Questions

Q. 1 Explain the difference of UDP and TCP socket of server side especially when the client initiates the connection or request to the server. **(Dec. 03, 10 Marks)**

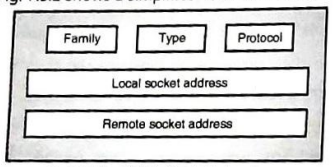
Q. 2 Explain the terms : Socket. **(Dec. 06, Dec. 07, May 09, 2 Marks)**

Q. 3 Explain with example : Socket. **(May 08, 10 Marks)**

- The socket interface was originally based on UNIX.
- It defines a set of system calls or procedure.
- The communication structure that we need in socket programming is called as a **socket.**
- A socket acts as an end point.
- Two processes can communicate if and only if both of them have a socket at their ends.

Socket structure :

- Fig. 7.5.1 shows a simplified socket structure.



(G-601) **Fig. 7.5.1 : Socket structure**

- Various fields in the socket structure are as follows :

1. **Family :** This field is used for defining the protocol group such as IPv4 or IPv6, UNIX domain protocol etc.
2. **Type :** This field is used for defining the type of socket such as stream socket, packet socket or raw socket.
3. **Protocol :** This field is usually set to zero for TCP and UDP.
4. **Local socket address :** It is used for defining the local socket address. This address is a combination of local IP address and the port address of the local application program.
5. **Remote socket address :** It is used for defining the remote socket address which is a combination of remote IP address and the port address of the remote application program.

7.5.1 Socket Types :

- There are three types of sockets :
  1. The stream socket
  2. The packet socket
  3. The raw socket
- All these sockets can be used in TCP/IP environment. Let us discuss them one by one.

1. Stream socket :

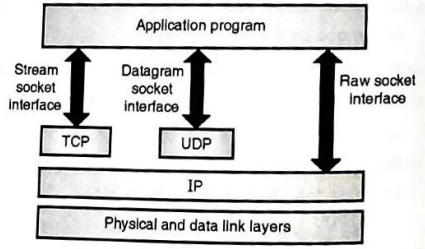
- This is designed for the connection oriented protocol such as TCP.
- The TCP uses a pair of stream sockets one each on either ends for connecting one application program to the other across the Internet.

2. Datagram socket :

- This type of socket is designed for the connectionless protocol such as UDP.
- UDP uses a pair of datagram sockets for sending a message from one application program to another across the Internet.

3. Raw socket :

- Raw sockets are designed for the protocols like ICMP or OSPF, because these protocols do not use either stream packets or datagram sockets.
- Fig. 7.5.2 shows the three types of socket types.



(G-602) **Fig. 7.5.2 : Type of sockets**

7.5.2 Berkeley Sockets:

MU : Dec. 09, Dec. 11, May 13, Dec. 13, May 15, May 19

University Questions

Q. 1 Write short notes on : Berkeley socket. **(Dec. 09, Dec. 11, May 13, 6 Marks)**

Q. 2 Show the usage of the different socket programming primitives used for establishing a connection between client and server. **(Dec. 13, 10 Marks)**

Q. 3 Write short note on Berkeley Socket. **(May 15, May 19, 5 Marks)**

- Table 7.5.1 lists various transport primitives used in Berkeley UNIX for TCP.

Table 7.5.1 : Various transport primitives

Sr. No.	Primitive	Meaning
1.	SOCKET	Create a new communication end point.
2.	BIND	Provide a local address to a socket
3.	LISTEN	Show willingness to accept connections
4.	ACCEPT	Block the caller as long as a connection attempt does not arrive
5.	CONNECT	Attempt to establish a connection
6.	SEND	Send data
7.	RECEIVE	Receive data
8.	CLOSE	Release the connection

- The first four primitives in the Table 7.5.1 are executed in the same order by the server.
- The SOCKET primitive creates a new end point and allocates table space for it within the transport entity.
- The newly created sockets do not have addresses. These are assigned using the **BIND** primitive.
- The **LISTEN** primitive allocates space to queue the incoming calls in case if several clients wish to connect at the same time.
- To block waiting for an incoming connection, the server executes an **ACCEPT** primitive.
- When a TPDU requesting for a connection arrives, the transport entity creates a new socket and returns a file descriptor for it.
- These were the primitives corresponding to server side. Now let us consider the client side.
- On the client side also a socket needs to be created first using the SOCKET primitive, however the BIND is not required.
- The **CONNECT** primitive blocks the caller and initiates the connection process.
- When it completes (which is indicated by an appropriate TPDU received from the server), the client process is unblocked and the connection is established.
- After this both the sides can use **SEND** and **RECEIVE** primitives to send and receive data.
- In order to release the connection, both sides have to execute a CLOSE primitive.

Steps followed for Socket Programming :

- The steps followed for the socket programming are as follows :

Server side :

1. Server creates a socket and checks for errors using SOCKET.
2. Assign address to the newly created socket using BIND.
3. Use the LISTEN to allocate space for the queue which is used for the incoming calls.
4. Execute an ACCEPT for blocking the waiting incoming connections.

Client side :

1. Create a socket using SOCKET.
2. Use CONNECT to initiate connection process.
3. Establish the connection.

7.6 Elements of Transport Protocols :

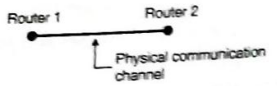
MU : Dec. 14

University Questions

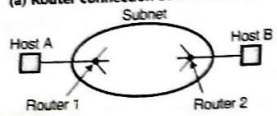
Q. 1 Explain the different elements of transport protocols. **(Dec. 14, 10 Marks)**

- In order to implement the transport layer services between the two transport entities, we have to use a **transport protocol.**
- The transport protocols have to deal with the following tasks :
  1. Error control
  2. Sequencing and
  3. Flow control
- The transport protocols are similar to the data link protocols in many ways but there are some dissimilarities as well.

- At the data link layer two router communicate directly via a physical channel as shown in Fig. 7.6.1(a), whereas at the transport layer the physical channel is replaced by the entire subnet as shown in Fig. 7.6.1(b).



(a) Router connection at data link layer



(b) Router connection at transport layer

(G-603) Fig. 7.6.1

- The difference between data link and transport layer communication is as given in Table 7.6.1.

Table 7.6.1 : Difference between data link and transport layer

Sr. No.	Data link layer	Transport layer
1.	Communication is through a physical channel.	Communication is through a subnet.
2.	It is not necessary to specify the destination router.	Explicit addressing of destination is essential.
3.	Establishing a connection is simple.	Initial connection establishment is more complicated.
4.	No storage capacity.	There is some storage capacity in the subnet.
5.	No additional delay.	Delay is introduced due to the storing capacity of subnets.
6.	Different approaches are used for buffering and flow control by the two layers.	

Elements of transport protocols :

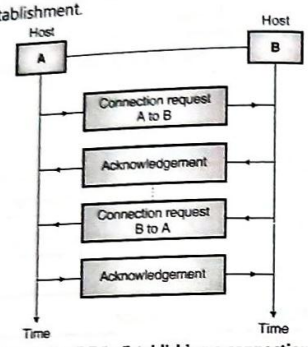
- Following are some of the important elements of transport protocols :
  1. Addressing
  2. Establishing a connection
  3. Releasing a connection
  4. Flow control and buffering
  5. Multiplexing
  6. Crash recovery

7.7 Connection Management :

- In a connection oriented service, a connection is established between source and destination. Then the data is transferred and at the end the connection is released.

7.7.1 Connection Establishment :

- Refer Fig. 7.7.1 to understand the connection establishment.



(G-604) Fig. 7.7.1 : Establishing a connection

- Following steps are taken to establish a connection :
  1. Host A sends a connection request packet to host B. This contains the initialisation information about data from A to B.
  2. Host B sends the packet of acknowledgement to confirm that it has received the request from A.
  3. Host B sends a connection request to A along with the initialisation information about traffic from B to A.
  4. Host A sends a packet of acknowledgement to confirm that it has received the request from B. It is possible to merge the steps 2 and 3.
- Note that each connection request must have a sequence number which is helpful in recovering from the loss or duplication of the packets.
- For the same reason, each acknowledgement also should have an acknowledgement number.
- The first sequence number in each direction should be random for each connection established. This is to ensure that a sender cannot create more than one connection which starts with the same sequence number (e.g. 2).

- This is important in recognizing the duplicate packets.
- Since a sequence number is required for each connection, the receiver has to keep the history of sequence numbers for each remote host for a specific amount of time but not indefinately.

Problems :

- Establishing a connection sounds easy. But actually it is a very tricky job. The problem occurs when the network can lose store and duplicate packets.
- The problems can be elaborated as follows :
  1. Due to congestion on a subnet, the acknowledgements do not get back in time from receiver to sender. So re-transmission of each packet takes place.
  2. If the subnet uses datagrams inside and every packet travels on a different route, then some of the packets might get stuck in a traffic jam and take a long time to arrive.
  3. The same connection getting re-established due to duplication of packet.
- So the crux of the problem is existence of delayed duplicate packets.

Remedy :

- The solution to this problem is to kill off the aged packets that are still wandering on the network.
- We should ensure that no packet lives longer than some predecided time.
- The packet lifetime can be restricted by using one of the following techniques :
  1. Restricted subnet design.
  2. Putting a hop counter in each packet.
  3. Time stamping each packet.

7.7.2 Three Way Handshake Technique :

MU : May 10, May 11, May 12, Dec. 13, May 15, Dec. 19

University Questions

Q. 1 Explain three way handshake technique in TCP. **(May 10, May 11, 10 Marks)**

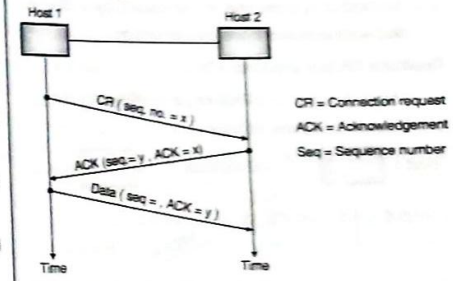
Q. 2 Explain the three protocol scenarios for establishing a connection using a 3-way handshake in TCP. **(May 12, 10 Marks)**

Q. 3 Show the different protocol scenarios for establishing a connection using 3-way handshake in the transport layer. **(Dec. 13, 10 Marks)**

Q. 4 Explain three way handshake technique in TCP. **(May 15, 10 Marks)**

Q. 5 Illustrate TCP three way handshake techniques in TCP connection establishment. **(Dec. 19, 10 Marks)**

- The delayed duplicate packet problem can be solved by using a technique called three way handshake.
- The principle of three way handshake is shown in Fig. 7.7.2(a).



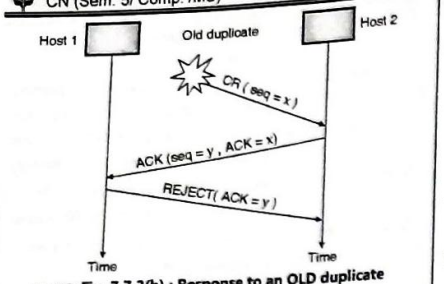
(G-605) Fig. 7.7.2(a) : Three way handshake technique

Normal operation :

1. Host 1 chooses a sequence number x and sends a TPDU containing the connection request (CR) TPDU to host 2.
2. Host 2 replies with a connection accepted TPDU to acknowledge x and to announce its own sequence number y.
3. Host 1 acknowledges host 2 and sends the first data TPDU to host 2.

Operation in the abnormal circumstances :

- Now let us see how the three way handshake works in presence of delayed duplicate control TPDUs.
- Refer Fig. 7.7.2(b). The first TPDU is a delayed duplicate CONNECTION REQUEST from an old connection. The HOST 1 does not know about it.

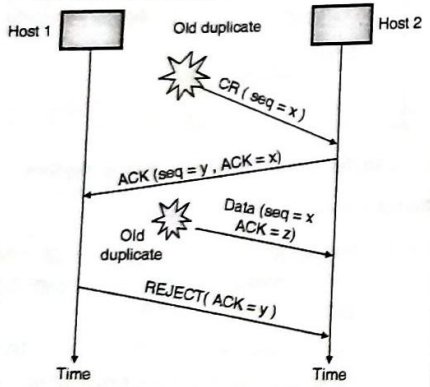


(G-606) Fig. 7.7.2(b) : Response to an OLD duplicate

- Host 2 receives this TPDU and sends to host 1, a connection accepted TPDU.
- But host 1 is not trying to establish any connection so it sends a REJECT alongwith ACK = y.
- So host 2 realizes that it was fooled by a delayed duplicate and abandons the connection.

Duplicate CR and duplicate ACK :

- This is another abnormal situation. Refer Fig. 7.7.2(c) to understand this situation.



(G-607) Fig. 7.7.2(c) : Duplicate CR and duplicate ACK

- *This is the worst case in which delayed duplicates of both connection request (CR) and acknowledgement (ACK) are making rounds in the subnet.*
- *Host 2 gets a delayed duplicate CR and it replies to it by sending ACK. Note that host 2 has proposed a connection with a sequence number y.*
- *When the second delayed TPDU (duplicate) arrives at host 2 it understands that z has been acknowledged and not y. So it understands that this too is an OLD duplicate.*

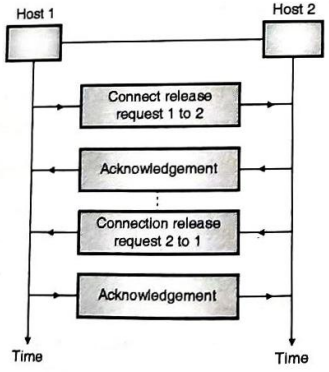
7.7.3 Connection Release :

- Any one of the two parties involved in data exchange can close the connection.
- But the problem is that, when connection is terminated from one end, the other party can continue to send data in the other direction.
- Hence, to ensure a proper connection release one has to follow the steps given below.

Procedure to release a connection :

- Refer Fig. 7.7.3 to understand this procedure.

1. Host 1 sends a connection release request to host 2.
2. Host 2 sends an acknowledgement to confirm the release request of host 1.
3. After this the connection is closed in one direction (no data from host 1 to host 2) but host 2 can continue to send data to host 1.
4. When host 2 finishes sending his data, it sends a connection release request to host 1.
5. Host 1 acknowledges (confirms) the request made by host 2 and the connection is released from both ends.



(G-608)Fig. 7.7.3 : Connection release

- Releasing a connection is easier than establishing it. There are two styles of releasing a connection :

Types of connecting release :

1. Asymmetric release and
2. Symmetric release

1. **Asymmetric release :**
   - In asymmetric release, when one party stops communicating the connection is broken. It is an abrupt release and it may lead to loss of data.
2. **Symmetric release :**
   - Symmetric release treats the connection as two separate unidirectional connections and in order to release the connection each side must release the connection on its side.
   - Symmetric release is as shown in Fig. 7.7.3 and there is no loss of data with the symmetrical release.

7.7.4 The Internet Transport Protocols (TCP and UDP) :

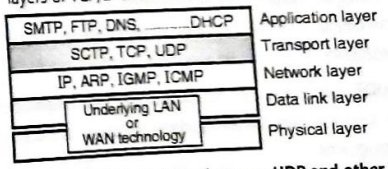
- The Internet has two main protocols in the transport layer.
- One of them is connection oriented and the other one supports the connectionless service.
- TCP (Transmission Control Protocol) is a connection oriented protocol and UDP (User's Data Protocol) is the connectionless protocol.
- UDP is basically just IP with an additional short header.

7.8 User Datagram Protocol (UDP) :

- The User Datagram Protocol is a very simple protocol. It adds little to the basic functionality of IP. Like IP, it is an unreliable, connectionless protocol.
- You do not need to establish a connection with a host before exchanging data with it using UDP, and there is no mechanism for ensuring that data sent is received.
- A unit of data sent using UDP is called a Datagram. UDP adds four 16-bit header fields (8 bytes) to whatever data is sent.
- These fields are : a length field, a checksum field, and source and destination port numbers. "Port number", in this context, represents a software port, not a hardware port.
- The concept of port numbers is common to both UDP and TCP.
- The port numbers identify which protocol module sent (or is to receive) the data.
- Most protocols have standard ports that are generally used for this.
- For example, the Telnet protocol generally uses port 23. The Simple Mail Transfer. Protocol (SMTP) uses port 25.
- The use of standard port numbers makes it possible for clients to communicate with a server without first having to establish which port to use.
- The port number and the protocol field in the IP header duplicate each other to some extent, though the protocol field is not available to the higher-level protocols. IP uses the protocol field to determine whether data should be passed to the UDP or TCP module.
- UDP or TCP use the port number to determine which application-layer protocol should receive the data.
- Although UDP isn't reliable, it is still a preferred choice for many applications.
- It is used in real-time applications like Net audio and video where, if data is lost, it's better to do without it than send it again out of sequence.
- It is also used by protocols like the Simple Network Management Protocol (SNMP).

Relationship with other protocols :

- The relationship of UDP with the other protocols and layers of TCP/IP suite is as shown in Fig. 7.8.1.



(G-2019) Fig. 7.8.1 : Relation between UDP and other protocols

- As shown, UDP is located between IP and application layer. It therefore works as an intermediary between application program and the network layer.

7.8.1 Responsibilities of UDP :

- Being a transport layer protocol, the UDP has the following responsibilities :

1. To create a process to process communication, UDP uses port numbers to accomplish this.

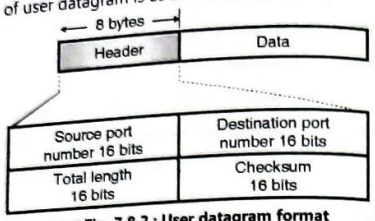
2. To provide control mechanisms at the transport layer, UDP does not provide flow control or acknowledgements. It provides error detection. The erroneous packet is discarded.
3. UDP does not add anything to the services of IP except for providing process to process communication.

7.8.2 Advantages of UDP :

- UDP, despite all its simplicity and powerlessness is still used because it offers the following advantages :
  1. UDP has minimum overheads.
  2. UDP can be easily used if the sending process is not too bothered about reliability.
  3. UDP reduces interaction between sender and receiver.

7.8.3 User Datagram :

- User Datagram Protocol (UDP) provides a connectionless packet service that offers unreliable 'best effort' delivery.
- This means that the arrival of packets is not guaranteed, nor is the correct sequencing of delivered packets.
- Applications that do not require an acknowledgement of receipt of data, for example, audio or video broadcasting uses UDP.
- UDP is also used by applications that typically transmit small amounts of data at one time, for example, the Simple Network Management Protocol (SNMP).
- UDP provides a mechanism that application programs use to send data to other application programs.
- UDP provides protocol port numbers used to distinguish between multiple programs executing on a single device.
- That is, in addition to the data sent, each UDP message contains both a destination port number and a source port number.
- This makes it possible for the UDP software at the destination to deliver the message to the correct application program, and for the application program to send a reply.
- UDP packets are called as **user datagrams.**
- They have a fixed-size header of 8-bytes. The format of user datagram is as shown in Fig. 7.8.2.



(G-624)**Fig. 7.8.2 : User datagram format**

- The UDP header is divided into the following four 16-bit fields :
  1. Source port
  2. Destination port
  3. Total length
  4. Checksum.

Source Port Number :

- Source port is an optional field, when meaningful, it indicates the port of the sending process, and may be assumed to be the port to which a reply should be addressed in the absence of any other information.
- If not used, a value of zero is inserted.
- This is a 16 bit field. That means the port numbers can range from 0 to 65,535.
- If the source host is a client, means if a client is sending a request using UDP, then generally a **ephemeral (temporary)** port number is requested by the process and chosen by the UDP.
- If the source host is a server that means if a server is sending a response message, mostly the **well known port** number is used.

Destination Port Number :

- The destination port number also is a 16 bit number and this port number is used by the process running on the destination host.
- If the destination host is a server that means if a client is sending a request to it, then a **well known port** number is used in most cases.
- However if the destination host is a client than means if a server is sending its response to it, then the chosen port number is generally an **ephemeral port** number.

Length :

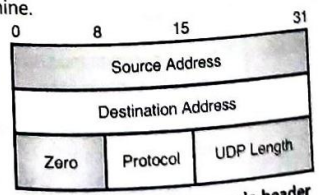
- It is also a 16 bit field which is used for defining the total length of the UDP datagram including header as well as data.
- Due to 16 bit length it can define a total length of the datagram upto 65,535 bytes.
- However practically the total length of a UDP datagram is much smaller than 65,535 bytes.
- This is because the UDP datagram is to be stored in an IP datagram which itself has a length of 65,535 bytes.
- The **length** field in the UDP datagram is actually not necessary, because this UDP datagram is actually encapsulated in an IP datagram and the IP datagram has its own length field.
- So without using the length field in UDP datagram, we can obtain the length of the UDP datagram as follows :
  UDP length = IP length – IP header length
- Note that while delivering the UDP datagram to UDP layer, the IP software drops the IP header.

UDP Checksum :

- This is used to verify the integrity (i.e. to detect errors) of the UDP header.
- The checksum is performed on a "pseudo header" consisting of information obtained from the IP header (source and destination address) as well as the UDP header.

7.8.4 UDP Pseudo Header :

- The purpose of using a pseudo-header is to verify that the UDP packet has reached its correct destination.
- The correct destination consists of a specific machine and a specific protocol port number within that machine.



(G-625) **Fig. 7.8.3 : UDP pseudo header**

- The UDP header itself specifies only the protocol port number.
- Thus, to verify the destination, UDP on the sending machine computes a checksum that covers the destination IP address as well as the UDP packet.
- At the ultimate destination, UDP software verifies the checksum using the destination IP address obtained from the header of the IP packet that carried the UDP message.
- If the checksum agrees, then it must be true that the packet has reached the intended destination host as well as the correct protocol port within that host.

User Interface :

- A user interface should allow the creation of new receive ports, receive operations on the receive ports that return the data octets and an indication of source port and source address, and an operation that allows a datagram to be sent, specifying the data, source and destination ports and addresses to be sent.

IP Interface :

- The UDP module must be able to determine the source and destination Internet addresses and the protocol field from the Internet header.
- One possible UDP/IP interface would return the whole Internet datagram including the entire Internet header in response to a receive operation.
- Such an interface would also allow the UDP to pass a full Internet datagram complete with header to the IP to send.
- The IP would verify certain fields for consistency and compute the Internet header checksum.

Protocol Application :

- The major uses of this protocol are the Internet Name Server, and the Trivial File Transfer.

Protocol Number :

- This is protocol 17 (21 octal) when used in the Internet Protocol.

Ex. 7.8.1 : The dump of a UDP header in hexadecimal format is as follows :

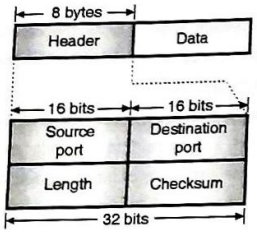
B C 8 2 0 0 0 D 0 0 2 B 0 0 1 D

Obtain the following from it :

1. Source port number
2. Destination port number
3. Total length
4. Length of the data.
5. Packet direction.
6. Name of client process.

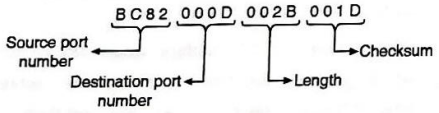
Soln. :

- The standard format of UDP header has been shown in Fig. P. 7.8.1.



(G-2020) Fig. P. 7.8.1 : UDP header format

- Therefore we can split the given UDP header in 4 equal parts as follows :



(G-2021)

1. Source port number = $(BC82)_{16}$ **...Ans.**
2. Destination port number = $(000D)_{16}$ **...Ans.**
3. Total length of UDP packet = $(002B)_{16}$
   = $(43)_{10}$ bytes **...Ans.**
4. Length of data = Total length – Length of the header.
   = 43 – 8 = 35 bytes **...Ans.**
5. Destination port number is $(000D)_{16} = (13)_{10}$
   - *It is a well known port. Hence the direction of UDP packet travel is from client to server.*
6. *The client process can be obtained from Table 7.9.1 which shows that for well known port number 13, the corresponding client process is "Daytime".*

7.9 UDP Services :

- In this section we are going to discuss the following important services provided by the UDP :
  1. Process to process communication.
  2. Connectionless services.
  3. Flow control.
  4. Error control.
  5. Checksum.
  6. Congestion control.
  7. Encapsulation and decapsulation.
  8. Queuing.
  9. Multiplexing and demultiplexing.

7.9.1 Process to Process Communication :

- We have already discussed the process to process communication in a general sense, earlier in this chapter.
- UDP also does it with the help of sockets which is a combination of IP address and port numbers.
- Table 7.9.1 shows different port numbers used by UDP.
- Some of these ports can be used by UDP as well as TCP.

Table 7.9.1 : Well known ports used with UDP

Port	Protocol	Description
7	Echo	The received datagram is echoed back to sender.
9	Discard	Any received datagram is discarded.
11	Users	Active users.
13	Daytime	Return the day and the current time.
17	Quote	Return the quote of the day.
19	Chargen	To return a string of characters.
53	Nameserver	Domain Name Service (DNS).
67	BOOT PS	This is the server port to download the bootstrap information.
68	BOOT PC	This is the client port to download bootstrap information.
69	TFTP	Trivial File Transport Protocol.
111	RPC	Remote Procedure Call.
123	NTP	Network Time Protocol.
161	SNMP	Simple Network Management Protocol.
162	SNMP	Simple Network Management Protocol (Trap).

7.9.2 Connectionless Services :

- As UDP is a connectionless, unreliable protocol, each user datagram sent using UDP is an independent datagram.
- Different user datagrams sent by the UDP have absolutely no relationship between them.
- This is true even for those datagrams which are originating from the same process and being sent to the same destination.
- The user datagrams do not have any number.
- Also the connection establishment and release are not at all required.
- So each datagram is free to travel any path.
- Only those processes which are sending very short messages can successfully use the UDP.

7.9.3 Flow and Error Control :

- Being a connectionless protocol, UDP is a simple, unreliable protocol.
- It does not provide any flow control, hence the receiver can overflow with incoming messages.
- UDP does not support any other error control mechanism, except for the checksum.
- There are no acknowledgements sent from destination to sender.
- Hence the sender does not know if the message has reached, lost or duplicated.
- If the receiver detects any error using the checksum, then that particular datagram is discarded.

7.9.4 Checksum :

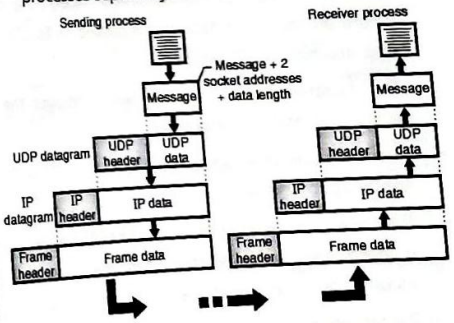
- The calculation of checksum for UDP is different than that for IP.
- In UDP the checksum is calculated by considering the following three sections :
  1. A pseudoheader
  2. The UDP header.
  3. The data coming from the application layer.
- The checksum in UDP is **optional.**
- That means the sender can make a decision of not calculating the checksum.
- If so, then the checksum field is filled with all zeros before sending the UDP packet.
- In case if the calculated checksum is all zeros (when the sender decides to send checksum) then an all 1 checksum is sent.
- This solution works without any problem because, a checksum will never have an all 1 value.

7.9.5 Congestion Control :

- UDP does not provide any congestion control.
- It assumes that the UDP packets being small, will not create any congestion.
- But this assumption may not always be correct.

7.9.6 Encapsulation and Decapsulation :

- The UDP encapsulates and decapsulates messages in an IP datagram in order to exchange the message between two communicating processes.
- This is as shown in Fig. 7.9.1. We will discuss the two processes separately.



(a) Encapsulation **(b) Decapsulation**

(G-2022) Fig. 7.9.1

Encapsulation :

- Refer Fig. 7.9.1(a). The message produced by a process is to be sent with the help of UDP.

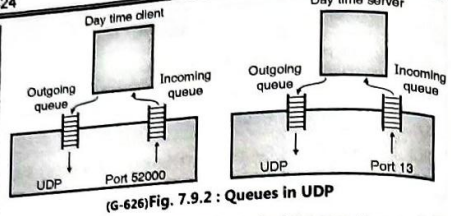
- The process passes the message and two socket addresses alongwith the length of data to UDP.
- UDP receives this data and adds the UDP header to it as shown.
- This is called as UDP datagram which is passed to IP with the socket address.
- IP adds its own header to UDP datagram as shown. It enters value 17 into the protocol field.
- This is an indication that UDP is being used. The IP datagram is then passed on to the data link layer.
- The DLL adds its own header and possibly a trailer to create a frame and sends it to the physical layer.
- Finally the physical layer converts these bits into electrical or optical signals and sends them to the destination machine.

Decapsulation :

- Refer Fig. 7.9.1(b) for understanding of the decapsulation process.
- The encoded message arrives at the destination physical layer where it decodes the electrical/optical signals into bits and passes them to the DLL.
- The DLL checks the data using header and trailer.
- The header and trailer are discarded if no errors are found, and the datagram is passed to IP.
- The IP carries out its checking to find the errors and if none are found, the datagram is passed on to UDP, after dropping the IP header.
- The datagram from IP to UDP also contains the sender and receiver IP addresses.
- This entire user datagram is checked by the UDP with the help of checksum.
- If there is no error detected, then the UDP header is dropped and the application data plus senders socket address are handed over to the process.
- The process can use this senders socket address if it wants to respond to the message received.

7.9.7 Queuing :

- The queues in UDP are related with ports as shown in Fig. 7.9.2.



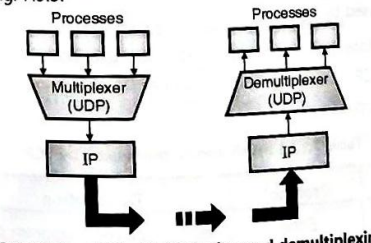
(G-626)**Fig. 7.9.2 : Queues in UDP**

- A process starts at the client site by requesting a port number from the operating system.
- In some implementations both incoming and outgoing queues are created in association with each process.
- Every process gets only one port number and hence it can create one outgoing and another incoming queue.
- The queues function only when the process is running. They are destroyed as soon as the process is terminated.
- The client process uses the source port number mentioned in the request to send message to its outgoing queue.
- UDP removes the queue messages one by one by adding the UDP header and delivers them to IP.
- If the outgoing queue overflows, then operating system tells that client process to wait before sending the next message.
- When the client receives a message, UDP checks if the incoming queue has been created or not.
- If the queue has been created, then the UDP sends the received datagram to the end of the queue.
- If the queue is not present then UDP will simply discard the user datagram.
- If the incoming queue overflows, then UDP discards the user datagram and arranges to send the port unavailable message to the server.
- The mechanism to create the server queue is different.
- The server creates the incoming and outgoing queues using its well known port as soon as it starts running.
- The queues exist as long as the server is running.

- When a message is received at the server, the UDP checks if the incoming queue has been created or not.
- If the queue is not present, the UDP discards the user datagram.
- If the queue is present then UDP sends the datagram at the end of the queue.
- If the incoming queue overflows, then UDP drops the user datagram and arranges to send the port unavailable message to the client.
- When the server wants to send a message to client it sends that message to the outgoing queue.
- These messages are then removed one by one after adding the UDP header. They are delivered to IP.
- If the outgoing queue overflows then the operating system will ask the server to wait before it sends the next message.

7.9.8 Multiplexing and Demultiplexing :

- We have discussed the general principle of multiplexing and demultiplexing in the transport layer.
- Now let us see how to apply the same principle to UDP. Imagine that a host is running a TCP/IP protocol suite and that there is only one UDP and a number of processes which would like to use the services of UDP.
- UDP handles such a situation by using the principle of multiplexing and demultiplexing as shown in Fig. 7.9.3.



(G-2023) **Fig. 7.9.3 : Multiplexing and demultiplexing**

Multiplexing :

- At the sending end, there are several processes that are interested in sending packets.
- But there is only one transport layer protocol (UDP or TCP).
- Thus it is a many processes-one transport layer protocol situation.
- Such a many-to-one relationship requires multiplexing.
- The UDP first accepts messages from different processes.
- These messages are separated from each other by their port numbers. Each process has a unique port number assigned to it.
- Then the UDP adds header and passes the packet to IP as shown in Fig. 7.9.3.

Demultiplexing :

- At the receiving end, the relationship is one as to many. So we need a demultiplexer.
- First the UDP layer receives datagrams from the IP.
- The UDP then checks for errors and drops the header to obtain the messages and delivers them to appropriate process based on the port number.

7.9.9 UDP Features :

- Some of the important features of UDP are as follows :

  UDP provides a connectionless service

  It does not provide error control

  It does not provide any congestion control

7.10 UDP Applications :

- Despite being connectionless, unreliable, no flow control, no error control, UDP is still preferred for some applications.
- This is because UDP has some advantages too. An application designer has to sometimes compromise between advantages and drawbacks to get the optimum.
- Some of the typical applications of UDP are as follows :

  1. UDP is suitable for the applications (processes) that have the following requirements :

     (a) A simple response to request is to be made.

     (b) Flow and error controls not essential.

     (c) Bulk data is not to be sent (like FTP).

2. UDP is used for RIP (Routing Information Protocol).
3. UDP is used for management processes such as SNMP.
4. UDP is suitable for the processes having inbuilt flow and error control mechanisms, such as TFTP.
5. UDP is suitable for the multicasting applications.
6. UDP is also used in the real time applications which do not tolerate the uneven delays.

7.11 Transmission Control Protocol (TCP) :

- The TCP provides reliable transmission of data in an IP environment. TCP corresponds to the transport layer (Layer 4) of the OSI reference model.
- Among the services TCP provides are stream data transfer, reliability, efficient flow control, full-duplex operation, and multiplexing.
- TCP is the layer 4 protocol in the TCP/IP suite and it is a very important and complicated protocol.
- TCP has been revised multiple times in last few decades.
- With stream data transfer, TCP delivers an unstructured stream of bytes identified by sequence numbers.
- This service benefits applications because they do not have to chop data into blocks before handing it off to TCP.
- Instead, TCP groups bytes into segments and passes them to IP for delivery.
- TCP offers reliability by providing connection-oriented, end-to-end reliable packet delivery through an internetwork.
- It does this by sequencing bytes with a forwarding acknowledgment number that indicates to the destination the next byte the source expects to receive.
- Bytes not acknowledged within a specified time period are retransmitted.
- The reliability mechanism of TCP allows devices to deal with lost, delayed, duplicate, or misread packets.
- A time-out mechanism allows devices to detect lost packets and request retransmission.
- TCP offers efficient flow control, which means that, when sending acknowledgments back to the source, the receiving TCP process indicates the highest sequence number that it can receive without overflowing its internal buffers.
- TCP supports a full-duplex operation means that TCP processes can both send and receive at the same time.
- Finally, TCP's multiplexing means that numerous simultaneous upper-layer conversations can be multiplexed over a single connection.

7.12 TCP Services :

- Following are some of the services offered by TCP to the processes at the application layer :
  1. Stream delivery service
  2. Sending and receiving buffers
  3. Bytes and segments
  4. Full duplex service
  5. Connection oriented service
  6. Reliable service.
  7. Process to process communication.

7.12.1 Process to Process Communication :

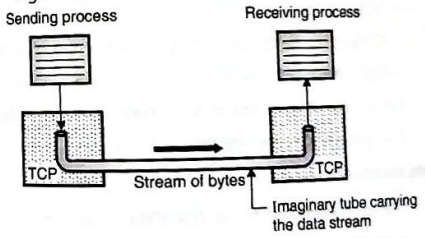
- The TCP uses port numbers a transport layer addresses.
- Table 7.12.1 shows some well known port numbers used by TCP.
- Note that if an application can use both UDP and TCP, the same port number is assigned to this application.

Table 7.12.1 : Well known ports used by TCP

Port	Protocol	Description
7	Echo	Sends received datagram back to sender
9	Discard	Discards any received packet
11	Users	Active users
13	Daytime	Sends the date and the time
17	Quote	Sends a quote of the day
19	Chargen	Sends a string character
20	FTP, Data	File Transfer protocol for data
21	FTP, Control	File Transfer protocol for control
23	TELNET	Terminal network
25	SMTP	Simple Mail Transfer Protocol
53	DNS	Domain Name server
67	BOOTP	Bootstrap Protocol
79	Finger	Finger
80	HTTP	Hypertext Transfer Protocol
111	RPC	Remote Procedure Call

7.12.2 Stream Delivery Service :

- TCP is a stream oriented protocol. The sending process delivers data in the form of a stream of bytes and the receiving process receives it in the same manner.
- TCP creates a working environment in such a way that the sending and receiving processes seem to be connected by an imaginary "tube" as shown in Fig. 7.12.1.

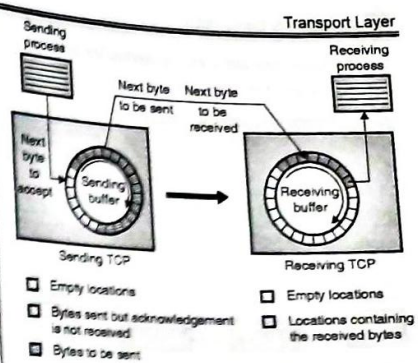


(G-621)Fig. 7.12.1 : Stream delivery service

- This is called as stream delivery service.

7.12.3 Sending and Receiving Buffers :

- The sending and receiving processes may not produce and receive data at the same speed.
- Hence TCP needs buffers for storage of data at both the ends.
- There are two types of buffers used in each direction :
  1. Sending buffer 2. Receiving buffer.
- A buffer can be implemented by using a circular array of 1 byte locations as shown in Fig. 7.12.2.



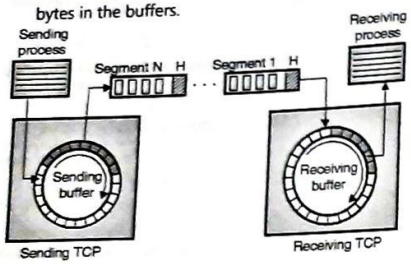
(G-622)Fig. 7.12.2 : Sending and receiving buffers

- Fig. 7.12.2 shows the direction of movement of data. The sending buffer has three types of locations :
  1. Empty locations
  2. Locations containing the bytes which have been sent but not acknowledged. These bytes are kept in the buffer till an acknowledgement is received.
  3. The locations containing the bytes to be sent by the sending TCP.
- In practice, the TCP may be able to send only a part of data which is to be sent, due to slowness of the receiving process or congestion in the network.
- The buffer at the receiver is divided into two parts :
  1. The part containing empty locations.
  2. The part containing the received bytes which can be consumed by the sending process.

7.12.4 Bytes and Segments :

- Buffering is used to handle the difference between the speed of data transmission and data consumption.
- But only buffering is not enough. We need one more step before sending the data.
- The IP layer, which provides service to TCP, has to send data in the form of packets instead of stream of bytes.
- At the transport layer, TCP groups a number of bytes to form a packet called a segment.
- A header is added to each segment for the purpose of exercising control.

- The segments are then inserted in an IP datagram and transmitted.
- The entire operation is transparent to the receiving process.
- The segments may be received out of order, lost or corrupted when it reaches the receiving end.
- Fig. 7.12.3 shows the creation of segments from the bytes in the buffers.



(G-623)Fig. 7.12.3

- The segments are not of the same size. Each segment can carry hundreds of bytes.

7.12.5 Full Duplex Service :

- TCP offers full duplex service where the data can flow in both the directions simultaneously.
- Each TCP will then have a sending buffer and receiving buffer.
- The TCP segments can travel in both the directions, therefore TCP provides a full duplex service.

7.12.6 Connection Oriented Service :

- TCP is a connection oriented protocol.
- When process – 1 wants to communicate (send and receive) with another process (process – 2), the sequence of operations is as follows :
  1. TCP of process – 1 informs TCP of process – 2 and create a connection between them.
  2. TCP of process – 1 and TCP of process – 2 exchange data in both the directions.
  3. After completing the data exchange, when buffers on both sides are empty, the two TCPs destroy their buffers to terminate the connection.
- The type of connection in TCP is not physical, it is virtual.
- The TCP segment is encapsulated in an IP datagram and these packets can be transmitted without following the sequence.
- These segments can get lost or corrupted and may have to be resent.
- Each segment may take a different path to reach the destination.

7.12.7 Reliable Service :

- TCP is a reliable transport protocol and not unreliable like UDP.
- Different acknowledgements are used by the receiver to convey sender the status of data.

7.13 Features of TCP :

- In order to provide the services mentioned in the previous section, TCP has a number of features as follows :

7.13.1 Numbering System :

- The TCP software keeps track of the segments being transmitted or received.
- However in the segment header there is no field for a segment number value.
- But there are fields called sequence number and the acknowledgement number.
- Note that these fields correspond to the byte number and not the segment number.

Byte numbers :

- TCP give numbers to all the data bytes which are transmitted.
- The numbering is independent of the direction of data travel.
- The numbering does not always start from 0, but it can start with a randomly generated number between 0 and $2^{32} - 1$.

Sequence number :

- After numbering the bytes, the TCP assigns a sequence number to each segment that is being transmitted.
- The sequence number for each segment is same as the number assigned to the first byte present in that segment.

Acknowledgement number :

- The TCP communication is duplex.
- So both the communicating processes can send and receive data at the same time.
- Each process will give numbers to the bytes with a different starting byte number.
- Each party also uses an acknowledgement number to confirm the reception of bytes.
- The acknowledgement number is cumulative i.e. the receiver takes the number of the last byte received, adds 1 to it and uses this sum as the acknowledgement number.

7.13.2 Flow Control :

- TCP provides flow control (UDP does not). The receiver will control the amount of data to be sent by the sender.
- This will avoid data overflow at the receiver. The TCP uses byte oriented flow control.

7.13.3 Error Control :

MU : Dec. 14

University Questions

Q. 1 Explain how TCP handles error control and flow control. **(Dec. 14, 10 Marks)**

- The error control mechanism is inbuilt for TCP. This allows TCP to provide a reliable service.
- The error control mechanism considers a segment as the unit of data for error correction however the byte oriented error control is provided.

7.13.4 Congestion Control :

- TCP takes the congestion in network into account. UDP does not do this.
- The amount of data sent by the sender depends on the following factors :
  1. The receivers decision (flow control).
  2. The network congestion.

Summary of TCP features :

1. TCP is a process-to-process protocol.
2. TCP uses port numbers.
3. It is a connection oriented protocol (creates a virtual connection).
4. It uses flow and error control mechanisms.
5. TCP is a reliable protocol.

7.14 The TCP Protocol : MU : Dec. 10, Dec. 11

University Questions

Q. 1 What is the function of TCP protocol ? Discuss its header format. **(Dec. 10, Dec. 11, 10 Marks)**

- Let us take a general overview of the TCP protocol.
- Every byte on a TCP connection has its own 32-bit sequence number.
- These numbers are used for both acknowledgement and for window mechanism.

Segments :

- The sending and receiving TCP entities exchange data in the form of segments.
- A segment consists of a fixed 20 byte header (plus and optional part) followed by zero or more data bytes.

Segment size :

- The segment size is decided by the TCP software. Two limits restrict the segment size as follows :
  1. Each segment including the TCP header, must fit in the 65535 byte IP payload.
  2. Each segment must fit in the **MTU (Maximum Transfer Unit).** Each network has a maximum transfer unit. Practically an MTU which is a few thousand bytes defines the upper limit on the segment size.

Fragmentation :

- If a segment is too large, then it should be broken into small segments. Using fragmentation by a router.
- Each new segment gets a new IP header. So the fragmentation by router will increase the overhead.

Timer :

- The basic protocol used by TCP entities is the sliding window protocol.
- A sender starts a timer as soon as a sender transmits a segment.
- When the segment is received by the destination, it sends back acknowledgement alongwith data if any.

- The acknowledgement number is equal to the next sequence number it expects to receive.
- If the timer at the sender goes out before the acknowledgement reaches back, it will **retransmit** that segment again.

Possible problems :

- As the segments can be fragmented, a part of the transmitted segment only may reach the destination with the remaining part lost.
- Segments can arrive out of order.
- Segments can get delayed so much that timer is out and unnecessary retransmission will take place.
- If a retransmitted segment takes a different route than the original segment is fragmented then the fragments of original and retransmitted segments can reach the destination in a sporadic way.
- So a careful administration is required to achieve reliable byte stream.
- There is a possibility of congestion or broken network along the path.
- TCP should be able to solve these problems in an efficient manner.

7.14.1 TCP Segment :

- The TCP segment as shown in Fig. 7.14.1 consists of two parts :

1. Header 2. Data

(G-1423)Fig. 7.14.1 : TCP segment

7.14.2 The TCP Segment Header :

MU : May 05, May 08, Dec. 08, Dec. 10, Dec. 11, May 19

University Questions

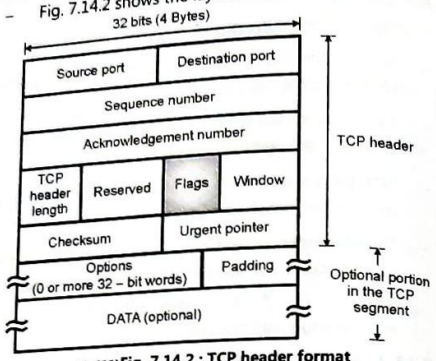
Q. 1 Draw and explain TCP header. **(May 05, 10 Marks)**

Q. 2 State different TCP flags. **(May 08, Dec. 08, 10 Marks)**

Q. 3 What is the function of TCP protocol ? Discuss its header format. **(Dec. 10, Dec. 11, 10 Marks)**

Q. 4 Describe TCP header with diagram. **(May 19, 10 Marks)**

- Fig. 7.14.2 shows the layout of a TCP segment.



(G-611)Fig. 7.14.2 : TCP header format

- Every segment begins with a 20 byte fixed format header.
- The fixed header may be followed by header options.
- After the options, if any, upto 65535 – 20 – 20 = 65495 data bytes may follow.
- Note that the first 20 bytes correspond to the IP header and the next 20 correspond to the TCP header.
- The TCP segment without data are used for sending the acknowledgements and control messages.

Source port :

- A 16-bit number identifying the application the TCP segment originated from within the sending host.
- The port numbers are divided into three ranges, well-known ports (0 through 1023), registered ports (1024 through 49,151) and private ports (49,152 through 65,535).
- Port assignments are used by TCP as an interface to the application layer.

Destination port :

- A 16-bit number identifying the application the TCP segment is destined for on a receiving host.
- Destination ports use the same port number assignments as those set aside for source ports.

Sequence number :

- A 32-bit number identifying the current position of the first data byte in the segment within the entire byte stream for the TCP connection.

- After reaching $2^{32} - 1$, this number will wrap around to 0.

Acknowledgement number :

- A 32-bit number identifying the next data byte the sender expects from the receiver.
- Therefore, the number will be one greater than the most recently received data byte.
- This field is only used when the ACK control bit is turned on.

Header length or offset :

- A 4-bit field that specifies the total TCP header length in 32-bit words (or in ultiples of 4 bytes if you prefer).
- Without options, a TCP header is always 20 bytes in length. The largest a TCP header may be is 60 bytes.
- This field is required because the size of the options field(s) cannot be determined in advance.
- Note that this field is called "data offset" in the official TCP standard, but header length is more commonly used.

Reserved :

- A 6-bit field currently unused and reserved for future use.

Control bits or flags :

1. **Urgent pointer (URG) :** If this bit field is set, the receiving TCP should interpret the urgent pointer field.
2. **Acknowledgement (ACK) :** If this bit field is set, the acknowledgement field described earlier is valid.
3. **Push function (PSH) :** If this bit field is set, the receiver should deliver this segment to the receiving application as soon as possible.
   - An example of its use may be to send a Control-BREAK request to an application, which can jump ahead of queued data.
4. **Reset the connection (RST) :** If this bit is present, it signals the receiver that the sender is aborting the connection and all queued data and allocated buffers for the connection can be freely relinquished.
5. **Synchronize (SYN) :** When present, this bit field signifies that sender is attempting to "synchronize" sequence numbers.
   - This bit is used during the initial stages of connection establishment between a sender and receiver.
6. **No more data from sender (FIN) :** If set, this bit field tells the receiver that the sender has reached the end of its byte stream for the current TCP connection.

Window :

- A 16-bit integer used by TCP for flow control in the form of a data transmission window size.
- This number tells the sender how much data the receiver is willing to accept.
- The maximum value for this field would limit the window size to 65,535 bytes, however a "window scale" option can be used to make use of even larger windows.

Checksum :

- A TCP sender computes a value based on the contents of the TCP header and data fields.
- This 16-bit value will be compared with the value the receiver generates using the same computation.
- If the values match, the receiver can be very confident that the segment arrived intact.

Urgent pointer :

- In certain circumstances, it may be necessary for a TCP sender to notify the receiver of urgent data that should be processed by the receiving application as soon as possible.
- This 16-bit field tells the receiver when the last byte of urgent data in the segment ends.

Options :

- In order to provide additional functionality, several optional parameters may be used between a TCP sender and receiver.
- Depending on the option(s) used, the length of this field will vary in size, but it cannot be larger than 40 bytes due to the size of the header length field (4 bits).
- The most common option is the Maximum Segment Size (MSS) option.
- A TCP receiver tells the TCP sender the maximum segment size it is willing to accept through the use of this option.

- Other options are often used for various flow control and congestion control techniques.

Padding :

- Because options may vary in size, it may be necessary to "pad" the TCP header with zeros so that the segment ends on a 32-bit word boundary as defined by the standard.

Data :

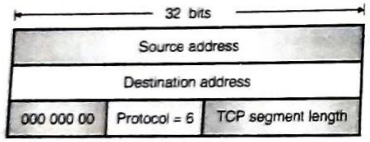
- Although not used in some circumstances (e.g. acknowledgement segments with no data in the reverse direction), this variable length field carries the application data from TCP sender to receiver.
- This field coupled with the TCP header fields constitutes a TCP segment.

7.14.3 Checksum : MU : Dec. 14

University Questions

Q. 1 Explain how TCP handles error control and flow control. **(Dec. 14, 10 Marks)**

- A checksum is provided to ensure extreme reliability. It checksums the header, the data and the conceptual pseudo header shown in Fig. 7.14.3.

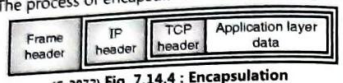


(G-612)Fig. 7.14.3 : The pseudo header included in the TCP checksum

- When the checksum is being computed, the TCP checksum field is set to zero, and the data field is padded out with an additional zero byte if its length is an odd number.
- Then all the 16 bit words are added in 1's complement and then 1's complement of the sum is taken to get the checksum.
- When a receiver performs the calculation on the entire segment including the checksum field, the result has to be zero.
- The pseudo header contains the 32 bit IP address of the source and destination machines, the protocol number for TCP i.e. 6 and the TCP segment length as shown in Fig. 7.14.3.

7.14.4 Encapsulation :

- The data coming from the application layer is encapsulated in a TCP segment.
- This TCP segment is then encapsulated in an IP datagram.
- The IP datagram is encapsulated in a frame at the data link layer.
- The process of encapsulation is shown in Fig. 7.14.4.



(G-2072) Fig. 7.14.4 : Encapsulation

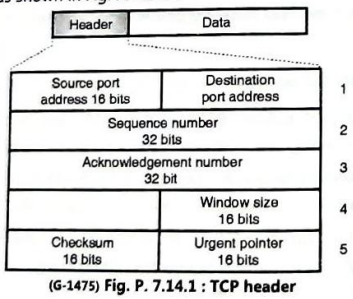
Ex. 7.14.1 : The following is a dump of TCP header in hexadecimal format

05320017 00000001 00000000 500207FF 00000000

1. What is the source port number ?
2. What is the sequence number ?
3. What is the acknowledgement number ?
4. What is the length of header ?
5. What is the window size ?

Soln. :

- The TCP segment format and its header format are as shown in Fig. P. 7.14.1.



(G-1475) Fig. P. 7.14.1 : TCP header

- The given TCP header consists of 5 words 32 bit each. We can identify them as follows :

(G-2364)

1. 0532 0017 — Destination port address (16 bit)
2. 0000 0001 — Sequence number (32 bit)
3. 0000 0000 — Acknowledgement number (32 bit)
4. 5002 07FF — Window size (16 bit)

Therefore :

1. Sequence number = 0000 0001 = 1
2. Destination port number = $(0\ 0\ 1\ 7)_H = 23$
3. Acknowledgement number = 0000 0000 = 0
4. Window size = 07FF = 2047 bytes

7.15 A TCP Connection :

- TCP is a connection oriented protocol. Such a protocol would establish a virtual path between the sender and the receiver.
- Multiple segments corresponding to the message are then sent over this virtual connection.
- As TCP is using the same single path for the entire path, it can use the same path for acknowledgements and retransmission of damaged or lost packets.
- While discussing the relation between TCP and IP we have seen how TCP uses the services of IP.
- TCP operates at a higher level than IP and the TCP connection is virtual and not physical.
- Though IP delivers the individual segments to the destination, the entire control on the connection is exercised by TCP.
- If a segment is lost or damaged, the TCP makes a decision of its retransmission, and IP does not know anything about it.
- The **three phases** in the connection oriented TCP transmission are as follows :
  1. Connection establishment
  2. Data transfer and
  3. Connection termination.

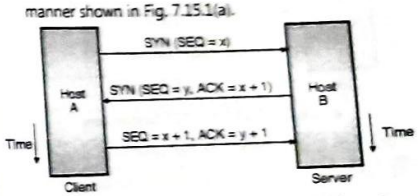
7.15.1 TCP Connection Establishment : MU : Dec. 03

University Questions

Q. 1 Show how TCP connection setup protects against the situation in :

Draw the space time diagram for protocol message exchange and explain how the protocol works. **(Dec. 03, 10 Marks)**

- To make the transport services reliable, TCP hosts must establish a connection-oriented session with one another.
- Connection establishment is performed by using a **three-way handshake** mechanism.
- A three-way handshake synchronizes both ends of a connection by allowing both sides to agree upon initial sequence numbers.
- This mechanism also guarantees that both sides are ready to transmit data and know that the other side is ready to transmit as well.
- This is necessary so that packets are not transmitted or re-transmitted during session establishment or after session termination.
- Each host randomly chooses a sequence number used to track bytes within the stream it is sending and receiving.
- Then, the three-way handshake proceeds in the manner shown in Fig. 7.15.1(a).



(G-613) Fig. 7.15.1(a) : TCP connection establishment (Three-way handshake)

- The requesting end (HOST A) sends a SYN segment specifying the port number of the server that the client wants to get connected to, and the client's initial sequence number (x).
- The server (HOST B) responds with its own SYN segment containing the server's initial sequence number (y).
- The server also acknowledges the client's SYN by acknowledging the client's SYN plus one (x + 1). A SYN consumes one sequence number.
- The client must acknowledge this SYN from the server by acknowledging the server's SYN plus one. (SEQ. = x + 1, ACK = y + 1).
- This is how a TCP connection is established.

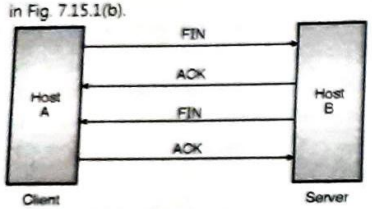
7.15.2 Connection Termination Protocol [Connection Release] :

MU : Dec. 03

University Questions

Q. 1 Show how TCP connection setup protects against the situation in :
Draw the space time diagram for protocol message exchange and explain how the protocol works.
(Dec. 03, 10 Marks)

- While it takes three segments to establish a connection, it takes four to terminate a connection.
- Since a TCP connection is full-duplex (that is, data flows in each direction independently of the other direction), the connection should be terminated in both the directions independently.
- The termination procedure in each direction is shown in Fig. 7.15.1(b).



(G-614) **Fig. 7.15.1(b) : TCP termination**

- The rule is that either side can send a FIN when it has finished sending data (FIN indicates finished).
- When a TCP program on a host receives a FIN, it informs the application that the other end has terminated the data flow.
- The receipt of a FIN only means there will be no more data flowing in that direction. A TCP can still send data after receiving a FIN.
- The end that first issues the close (e.g., sends the first FIN) performs the active close and the other end (that receives this FIN) performs the passive close.
- Now refer Fig. 7.15.1(b). When the server receives the FIN it sends back an ACK of the received sequence number plus one.
- A FIN consumes a sequence number, just like a SYN.
- At this point the server's TCP also delivers an end-of-file to the application (the discard server).
- The server then closes its connection and its TCP sends a FIN to the client.
- The client's TCP informs the application and sends an ACK to server by incrementing the received sequence number by one.
- Connections are normally initiated by the client, with the first SYN going from the client to the server.
- A client or server can actively close the connection (i.e. send the first FIN).
- But in practice generally the client determines when the connection should be terminated, since client processes are often driven by an interactive user, who enters something like quit to terminate.
- This is how the TCP connection is released.

7.15.3 TCP Connection Management :

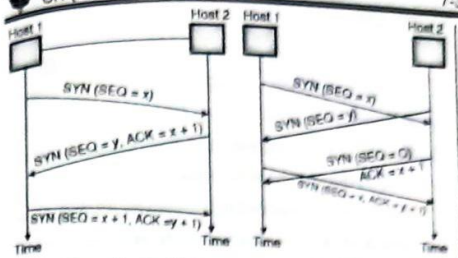
MU : Dec. 15, May 17, Dec. 17

University Questions

Q. 1 Write short notes on : TCP connection management. **(Dec. 15, 10 Marks)**

Q. 2 Explain with the help of suitable diagram TCP connection management and release. **(May 17, Dec. 17, 10 Marks)**

- Connections are established in TCP by following the three-way handshake technique.
- To establish a connection, one side, say the server, passively waits.
- It executes the LISTEN and ACCEPT primitives, to specify either a particular other side or nobody in particular.
- The other side (client) executes a connect primitive, with the IP and the port specified.
- The other information is the maximum TCP segment size, possible other options and optionally some user data (e.g. a password).
- The CONNECT primitive sends a TCP segment with the SYN bit on and the ACK bit off and waits for a response.
- The sequence of TCP segments sent in the normal case is shown in Fig. 7.15.2(a).



(a) Normal operation **(b) Call collision**

(G-615) **Fig. 7.15.2 : TCP connection management**

- When the segment sent by host – 1 reaches the destination i.e. host – 2 the receiving server checks to see if there is a process that has done a LISTEN on the port given in the destination port field.
- If not, it sends a reply with the RST bit on to reject the connection.
- Otherwise it gives the TCP segment to the listening process, which can accept or refuse (e.g. if it does not like the client) the connection.
- On acceptance a SYN is send, otherwise a RST.
- Note that a SYN segment occupies 1 byte of sequence space so it can be acknowledged unambiguously.

Call collision :

- If two hosts try to establish a connection simultaneously between the same two sockets then the events take place as shown in Fig. 7.15.2(b).
- Under such circumstances only one connection is established. Both the connections can not be established simultaneously because connections are identified by their end points.
- If the first set up results in a connection which is identified by (x, y) and second connection is also set up, then only one table entry will be made i.e. for (x, y).
- For the initial sequence number a clock based scheme is used, with a clock pulse coming after every 4 μsec.
- For ensuring an additional safety, when a host crashes, it may not reboot for 120 sec which is maximum packet lifetime.
- This is to make sure that no packets from previous connections are still alive and travelling around.

7.15.4 TCP Connection Release :

MU : May 17, Dec. 17

University Questions

Q. 1 Explain with the help of suitable diagram TCP connection management and release. **(May 17, Dec. 17, 10 Marks)**

- A TCP connection is actually a full duplex connection but to understand the connection release we will assume that it is a pair of simplex connections.
- We can then think that each simplex connection is getting terminated independently.
- Releasing a TCP connection is identical on both ends. Each side can send a TCP segment with the FIN bit set, meaning it has no more data to send.
- After receiving a FIN, the Acknowledge (ACK) signal is sent and that direction is shut down, but data may continue to flow indefinitely in the other direction.
- If the sender of FIN does not receive the ACK within 2 maximum packet lifetimes, it releases the connection. The receiver will eventually notice that it receives no more data and time-out as well.
- Normally four TCP segments are required to release a connection i.e. one FIN and one ACK in each direction.
- However the first ACK and second FIN can be combined in the same segment.

Connection reset :

- The connection reset in TCP can take place when TCP at one end done any one of the following :
  1. It may deny a connection request.
  2. It may abort the existing connection.
  3. It may terminate an idle i.e. non operating connection.
- TCP does all the three with the help of the RST (reset flag).

7.16 TCP State Transition Diagram :

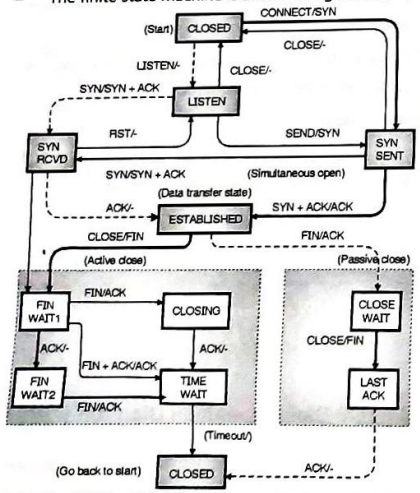
- The steps to be followed in TCP connection establishment and release can be represented using a finite state machine.

- The total eleven states in such a state machine are given in Table 7.16.1.

Table 7.16.1 : Different states in TCP finite state machine

State	Description
CLOSED	No connection is active or pending
LISTEN	The server is waiting for an incoming call
SYN RCVD	A connection request has arrived; wait for ACK
SYN SENT	The application has started to open a connection
ESTABLISHED	The normal data transfer state
FIN WAIT 1	The application has said it is finished
FIN WAIT 2	The other side has agreed to release
TIMED WAIT	Wait for all packets to die off
CLOSING	Both sides have tried to close simultaneously
CLOSE WAIT	The other side has initiated a release
LAST ACK	Wait for ack of FIN of last close

- In each of the 11 states shown in Table 7.16.1, some specific events are considered to be legal events.
- Corresponding to every legal event some action may be taken, but if some event other than the legal one happens, then error is reported.
- The finite state machine is shown in Fig. 7.16.1.



(G-696) Fig. 7.16.1 : TCP connection management final state machine

- Each connection is always in the CLOSED state initially. It comes out of this state when it does either the passive open (LISTEN) or an active open (CONNECT).
- A connection is established, if the other side does the opposite and the state becomes ESTABLISHED.
- When both the sides initiate a connection release the connection is terminated and the state returns to CLOSED state.

Various types of lines in the finite state machine drawing :

- Various types of lines are used in the finite state machine drawing of Fig. 7.16.1. They have different meanings as stated below :
  1. **Heavy solid lines :** These lines show a client actively connecting to a passive server.
  2. **Heavy dotted lines :** These lines are used for the server.
  3. **The light faced lines :** These are for unusual event sequences.
- Over each line in Fig. 7.16.1 we have written the event / action pair.
- The event can either be a user-initiated system call (CONNECT, LISTEN, SEND or CLOSE), a segment arrival (SYN, FIN, ACK or RST), or a time-out.
- For the TIMED WAIT state the event can only be a time-out of twice the maximum packet length. The action is the sending of a control segment (SYN, FIN or RST) or nothing.
- The time-outs to guard for lost packets (e.g. in the SYN SENT state) are not shown here.
- There are 11 states used in the TCP connection management finite state machine. Data can be send in the ESTABLISHED and the CLOSE WAIT states and received in the ESTABLISHED and FIN WAIT1 states.

Explanation :

- To understand the finite state machine of Fig. 7.16.1, first follow the path of a client i.e. the heavy solid line. After that follow the path of the server (the heavy dashed line).

7.17 Windows in TCP : MU : Dec. 10

University Questions

Q. 1 Discuss the window management in TCP transmission policy with a neat diagram. **(Dec. 10, 5 Marks)**

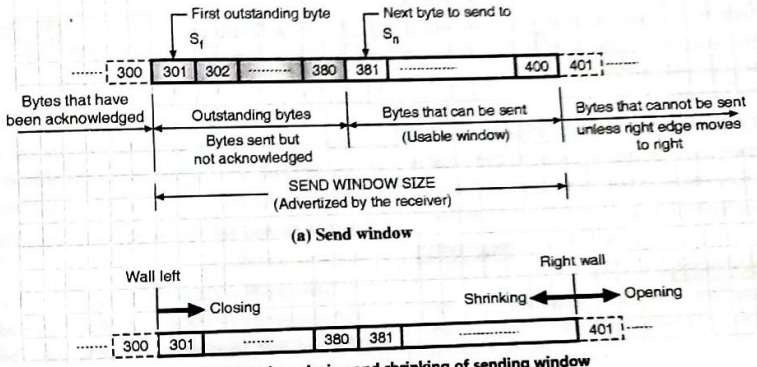
- In this section we will discuss the windows used in TCP. There are two types of windows used in TCP :
  1. Send window
  2. Receive window
- Send window is for sending data and receive window is for receiving data. Therefore there will be four windows in all for a two way communication.
- However in order to make the discussion simple, we will assume that the communication takes place only in one direction (client to server or the other way round).

7.17.1 Send Window : MU : Dec. 10

University Questions

Q. 1 Discuss the window management in TCP transmission policy with a neat diagram. **(Dec. 10, 5 Marks)**

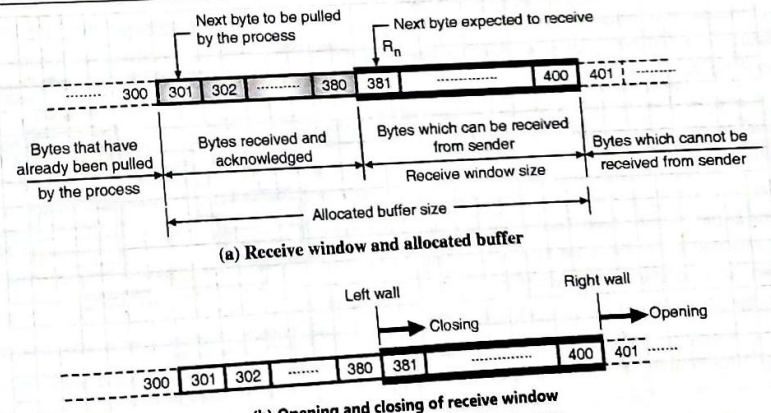
- Fig. 7.17.1 illustrates an example of send window. In reality the send window can have a size of thousands of bytes however for simplifying the discussion a 100 byte send window has been considered in Fig. 7.17.1.



(a) Send window

(b) Opening, closing and shrinking of sending window

(G-1800) Fig. 7.17.1 : Send window in TCP



(a) Receive window and allocated buffer

(b) Opening and closing of receive window

(G-1801) Fig. 7.17.2 : Receive window in TCP

- The size of send window is dependent on the receiver (flow control) as well as on the congestion control.
- There are three operations that can take place in the send window, namely : open, close and shrink.
- The send window in TCP is similar to that in selective repeat request (SR) with the following differences :
  1. The SR send window numbers pockets but TCP send window numbers bytes. In TCP the transmission takes place in the form of segments but the controlling parameters of windows are expressed in bytes.
  2. Actually TCP is capable of storing data received from the process and send it later on. But we will assume that the sending TCP sends the segments of data as soon as it is received from the process.
  3. TCP uses only one timer as compared to several timers used by the SR protocol. This timer in TCP is used for error control.

7.17.2 Receive Window :

MU : Dec. 10

University Questions

Q. 1 Discuss the window management in TCP transmission policy with a neat diagram. **(Dec. 10, 5 Marks)**

- The example of receive window has been shown in Fig. 7.17.2. In reality the receive window can have size of thousands of bytes however for simplification of discussion a 100 byte receive window has been shown in Fig. 7.17.2.
- The receive window in TCP is similar to that in selective repeat request (SR) with the following differences :
  1. The receiving process in TCP is allowed to pull data as per its own speed. That means in a part of allocated buffer there are bytes which have been received and acknowledged but waiting for the receiving process to pull them (see Fig. 7.17.2). The size of receive window is therefore always smaller than the allotted buffer size. The size of the receive window will decide the number of bytes a receiver can receive without causing the flow control problems. The receiver window size which is denoted by "rwnd" is expressed as follows :

rwnd = Buffer size – Number of acknowledged bytes to be pulled

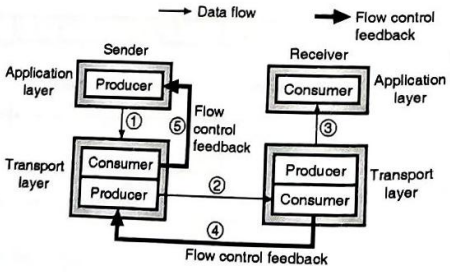
  2. The acknowledgements in SR define the uncorrupted received packets only. This is selective acknowledgement. However in TCP the mechanism of acknowledgement is called as **cumulative acknowledgement** in which the next expected byte to be received (R_n = 381 in Fig. 7.17.2) is announced. The new version of TCP uses both selective and cumulative mechanisms for acknowledgements.

7.18 Flow Control in TCP : MU : Dec. 14

University Questions

Q. 1 Explain how TCP handles error control and flow control. **(Dec. 14, 10 Marks)**

- The flow control is a technique used for controlling the data rate of the sender so that the receiver is not overwhelmed.
- In TCP the flow control has been kept separate from the error control.
- So when the flow control is being discussed, we will temporarily ignore the error control. i.e. we assume that the data transmission is taking place over an errorfree channel.
- Refer Fig. 7.18.1 which shows the data transfer taking place in only one direction from the sender to receiver.

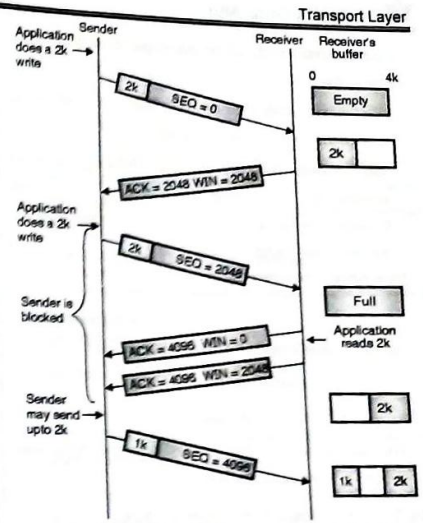


(G-1802) Fig. 7.18.1 : Data flow and flow control feedback in TCP

- We can apply the same principle to the bidirectional data transfer.
- Two different types of signals travel between the sending process and the receiving process in Fig. 7.18.1.
- They are data and flow control feedback signals.
- The data flow takes place from the sending process to the sending TCP (denoted by ①), then from sending TCP to receiving TCP (denoted by ②) and finally from receiving TCP to receiving process (denoted by ③).
- Thus flow of data takes place from sender to receiver. But the flow control feedback signals travel from the receiver to sender as shown.
- They flow from receiving TCP to sender TCP (denoted by ④) and from sending TCP to sending process (denoted by ⑤).
- Most TCP versions however, do not provide the flow control feedback facility.
- Instead the receiving process is allowed to pull data from receiving TCP whenever the receiving process becomes ready.
- Thus the receiving TCP controls the sending TCP (due to flow control feedback) and the sending TCP controls the sending process as far as the data rate of the sending process is concerned.
- Consider the flow control feedback path denoted by ⑤ in Fig. 7.18.1.
- This feedback is practically achieved by simply rejecting the data by sending TCP when its window is full.
- So now let us concentrate on the flow control feedback signal from receiving TCP to sending TCP, denoted by path ④ in Fig. 7.18.1. i.e. how does the receiving process control the sending TCP.

7.18.1 Opening and Closing Windows :

- In TCP the flow control is achieved by forcing the sender and receiver to adjust their window sizes.
- The size of the buffer for both sender and receiver will not be changed. It will remain fixed in size.
- Consider the receive window shown in Fig. 7.18.2.



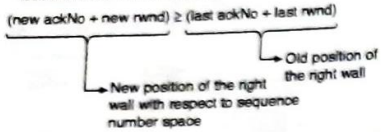
(G-616) Fig. 7.18.2 : Windows management in TCP

- This window closes by moving its left wall to the right in response to arrival of more bytes from the sender.
- The receive window of Fig. 7.18.2 will open by moving its right wall towards right when the receiver process pulls more bytes from the receiver buffer.
- The send window can open, close or shrink in order to exercise the flow control.
- All the three functions of the send window are controlled by the receiver.
- The send window closes by moving its left wall to the right (see Fig. 7.18.2) in response to a new acknowledgement from the receiver.
- The send window opens by moving its right wall to the right when the advertised receive window size (rwnd) by the receiver allows it to do so.
- The send window may shrink on occasion. It is assumed that this situation does not arise.

7.18.2 Shrinking of Windows :

- As we know, the receiver window does not shrink.
- However the send window can shrink in the event of the receiver defining a value of "rwnd" which results in the shrinking of windows.

- Some versions of TCP do not allow the send window to shrink.
- That means they do not allow the right wall of the send window to move to the left.
- The receiver can prevent the shrinking of send window by maintaining the following relationship between the last and new acknowledgement and the last and new "rwnd" values.



(G-1803)

- The above relationship shows that the right wall should not move to the left.

7.18.3 An Example of Flow Control :

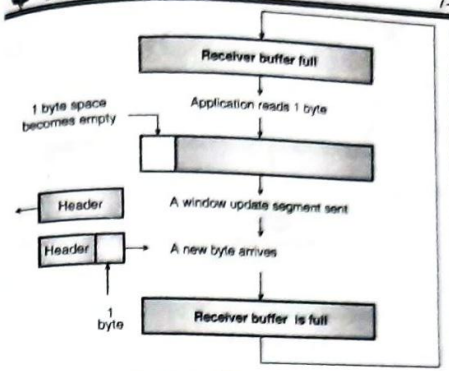
- Let us now see how the window policy is used in transmission policy of TCP protocol.
- Window management in TCP is normally decoupled from the acknowledgements that means acknowledgements are not connected to the TCP window management.
- To understand the window management, refer Fig. 7.18.2.

Explanation :

- Let the receiver in Fig. 7.18.2, has a 4 kbyte i.e. 4096 byte buffer space.
- The sender transmits a 2048 byte (2 kbyte) segment with a sequence number SEQ = 0.
- These bytes occupy half space of the receiver's buffer and the receiver will send back acknowledgement of this segment (ACK 2048, WIN = 2048).
- Here WIN = 2048 is the window which tells the sender that an empty buffer space of 2048 is available on the receiver side.
- Now the sender sends another 2k i.e. 2048 bytes segment (SEQ = 2048) which is acknowledged by the receiver (ACK = 4096, WIN = 0) which shows that window = 0 because the receiver buffer space is 0.
- ACK = 4096 indicates that the receiver has received 4096 bits successfully.
- The sender must now be blocked until the application process on the receiver removes some data from the buffer and some buffer space becomes available.
- As soon as the application on the receiver side reads 2k bytes, the buffer becomes partially empty and an acknowledgement with a window of 2k (ACK = 4096, WIN = 2048) is sent back to sender.
- Here WIN = 2048 indicates the empty buffer space on the receiver side.
- The sender may send upto 2 kbytes.
- When the window = 0, the sender should not normally send any segment.
- But under two exceptional conditions the sender will continue to send data even when it receives WIN = 0.
  1. First, urgent data may be send, e.g. to allow the user to kill the process running on the other machine.
  2. Second, the sender may send a 1-byte segment to make the receiver reannounce the next byte expected and the window size.
- This is used to prevent the possible confusion if a window announcement gets lost.
- Senders are not supposed to transmit data as soon as the data is obtained from an application.
- The receivers also are not supposed to send acknowledgements as soon as they receive it.
- This is done in order to reduce the usage of the system.
- One way to reduce the system usage is to use an algorithm called Nagle's algorithm is used.

7.18.4 Silly Window Syndrome :

- This is another problem that can degrade the TCP performance.
- This problem occurs when the sender transmit data in large blocks, but an interactive application on the receiver side reads data 1 byte at a time.
- To understand this problem, refer Fig. 7.18.3.



(G-617) **Fig. 7.18.3 : Silly window syndrome**

1. Initially the receiver's buffer is full so it send a window size 0 to block the sender.
2. But the interactive application reads one byte from the buffer. So one byte space becomes empty.
3. The receiving TCP sends a window update to the sender informing that it can send 1 byte.
4. The sender send 1-new byte.
5. The buffer is full again and the window size is 0. This process can continue forever. This is known as the silly window syndrome.

7.18.5 Nagle's Algorithm :

- The Nagle's algorithm is very simple. It takes into account the speed of transmission of the sender and the speed of the network which is transporting the data.
- The algorithm is as follows :
  1. The first piece of data received from the sending application program is send by the sending TCP even if it is only 1 byte.
  2. Once the first segment is sent, the sending TCP will wait and accumulate data in the output buffer until either the acknowledgement is received from the receiving TCP or sufficient data is accumulated to fill the maximum size segment.
  3. Step 2 is repeated for the remaining transmission.
- If the sending application program data rate is higher than the speed of data transporting network then the segments are larger (maximum size segments).
- On the other hand if the sending application program is slower than the data transport network, the segments will be smaller than the maximum segment size.

Clark's solution to silly window syndrome :

- Clark suggested a solution to silly window syndrome as follows.
- He suggested that the receiver should not send a window update for 1 byte.
- Instead the receiver must wait until it has a considerable amount of buffer space available and then send the window update.
- To be specific, the receiver should wait until it can handle the maximum window size it has advertised at the time of establishing a connection or its buffer is half empty, whichever is smaller.
- The sender can also help to improve the situation.
- It should not send tiny segments. Instead it must wait and send a full segment or at least one containing half of the receivers buffer size.

7.19 TCP Congestion Control :

MU : Dec. 07, May 08, May 09, Dec. 09, May 13

University Questions

- **Q. 1** What is congestion control ? How it is different from flow control ? Explain various congestion prevention techniques. **(Dec. 07, 10 Marks)**
- **Q. 2** What is congestion ? Explain how it can be avoided. **(May 08, 10 Marks)**
- **Q. 3** Explain how TCP controls congestion. **(May 09, 10 Marks)**
- **Q. 4** How TCP controls the congestion ? **(Dec. 09, 10 Marks)**
- **Q. 5** How TCP controls the congestion, explain in detail. **(May 13, 10 Marks).**

- We have already discussed the reasons of congestion in networks and the Internet is no exception.
- So there are congestions occurring on Internet too.
- The network layers detects the congestion by looking at the growing queues at the routers and tries to manage it by dropping packets.

- The network layer has to give feedback to the transport layer about the possible congestion because only then the transport layer can reduce the sender's data rate.
- In the Internet, TCP plays a major role in controlling congestion.
- A control law called AIMD (Additive Increase Multiplicative Decrease) can be used in response to binary congestion signals received from the network.
- According to this law, in response to congestion signals the transport protocol should converge to a fair and efficient bandwidth allocation.
- TCP congestion control is based on this approach using a **window** and with a loss of packet used as the binary signal to indicate congestion.

Principle of congestion control :

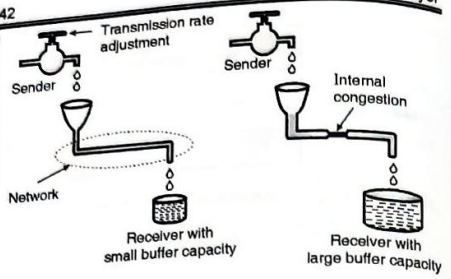
- The basic principle is do not inject a new packet into the network until an old one is delivered.
- TCP tries to do this by dynamically adjusting the window size.
- The steps followed in achieving the congestion control in TCP are as follows :

Step 1 : Detect the congestion :

- This is the first step in congestion control.
- Now-a-days packet loss due to transmission errors is very rare because the optical fiber links are being used.
- So most transmission time-outs (loss of packets) are due to congestions.
- So all the Internet TCP algorithms assume that time-outs are caused by congestion and so time outs can be used to detect the congestion.

Step 2 : Try to prevent congestion :

- After establishing a connection, a suitable window size is to be chosen.
- The receiver window size is based on its buffer capacity.
- If the sender adjusts its transmission rate according to this capacity as shown in Fig. 7.19.1(a), the congestion due to buffer overflow will never take place.



(a) No congestion **(b) Internal**

(G-618) Fig. 7.19.1 : Congestion

- Now consider Fig. 7.19.1(b). The sender is slow, the receiver has a large buffer capacity but the problem is low internal carrying capacity of the network.
- If the sender is too fast, the water will back up and some will be lost (loss of packets) and congestion will take place.

Conclusion :

- To prevent congestion TCP has to deal with two problems separately – receiver capacity and network capacity.

Solution :

- To deal with the two problems mentioned earlier each sender maintains two windows : the window the receiver has granted (which indicates the receiver capacity) and the **congestion window** (which indicates the network capacity).
- The first window that indicates the receiver capacity is called as the **flow control window.**
- The size of the congestion window is equal to the number of bytes the sender may have in the network at any time.
- Hence the corresponding **sending rate** is equal to the ratio of **congestion window size** and the **RTT** of the connection.
- TCP adjusts the size of window as per the AIMD rule.
- The **congestion window** is maintained in addition to the **flow control window** (Which specifies the number of bytes that the receiver can buffer).
- Both these windows are considered simultaneously.

- Both the windows indicate the number of bytes the sender may transmit and the number can be different.
- Therefore the number of bytes that may be sent by the sender is the minimum of the two windows.
- So the effective window is the minimum of what the sender and the receiver both think is all right.

Modern congestion control :

- Modern congestion control was added to TCP in 1988 through the efforts of Van Jacobson.
- In 1986 due to growing number of Internet users the first **congestion collapse** took place.
- As a response to this collapse Jacobson approximated an AMID congestion window and added it to the existing TCP.
- While doing so he made following two important considerations :

1. The rate at which the acknowledgements return to the sender is approximately equal to the rate at which packets can be sent over the slowest link in the path. This is the rate a sender wants to use to avoid congestion. This timing is known as **ACK clock** and it is an essential part of TCP. Using ACK clock TCP smoothes out traffic and avoids congestion.
2. The second consideration was that AMID rule will take a very long time to reach the desired operating point on fast networks if the congestion window is started from a small value. The start up time can be reduced by using a large initial window. But a too large starting window would cause congestion in slow or short links.

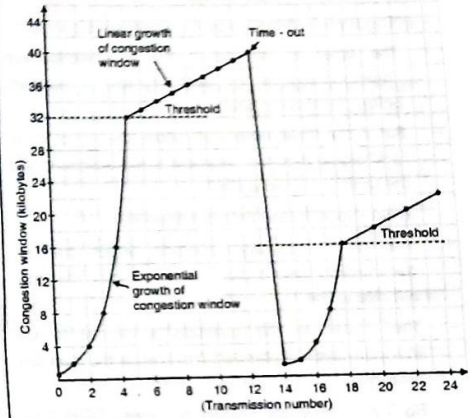
- Hence Jacobson mixed both linear and multiplicative increase in the window size in his solution to resolve congestion.
- This modified algorithm is known as the **slow start** algorithm.

7.19.1 Slow Start Algorithm : MU : Dec. 07

University Questions

Q. 1 What is congestion control ? How it is different from flow control ? Explain various congestion prevention techniques. **(Dec. 07, 10 Marks)**

1. After establishing a connection, the sender initialises the congestion window to the size which is equal to the maximum segment in use on the connection. It then sends one maximum segment.
2. If this segment is acknowledged by the receiver indicating no congestion, it adds bytes corresponding to one full segment to the congestion window. So now the congestion window size is equal to two maximum size segments. The sender then sends two segments.
3. As each of these segments is acknowledged indicating that there is no congestion, the size of congestion window is increased by one maximum segment size. This is shown in Fig. 7.19.2. This is the exponential growth of the congestion window size.



(G-619) Fig. 7.19.2

4. When the congestion window is of n segments, if all n segments are acknowledged before time-out takes place, the congestion window is increased by the byte count corresponding to n segments.
5. But there is a limit on the exponentially growing congestion window. The congestion window stops growing as soon as either the time-out occurs or the receiver's window size is reached.
6. If the congestion window can grow to 1024 (1 kbyte) byte, 2048 byte, but a burst of 4096 bytes gives a time-out then we have to set the congestion window at 2048 in order to avoid congestion.

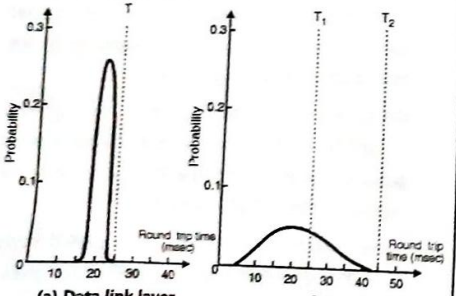
7. Once this is done, no data bursts longer than 2048 bytes will be sent by the sender even if receiver grants a wider window.
8. The name of this algorithm is slow algorithm and it is required to be supported by all the TCP implementations.

7.20 TCP Timer Management : MU : Dec. 18

University Questions

Q. 1 Explain the use of TCP timers in detail. **(Dec. 18, 10 Marks)**

- The TCP, at least conceptually uses more than one timers.
- But the most important of them is the **Re-transmission Timer (RTO).**
- This timer is started as soon as a segment is sent.
- The timer is stopped if the acknowledgement corresponding to the sent segment is received, before the timer expires.
- But if the timer times out before the arrival of an "ack" signal then that segment is re-transmitted and the timer is started again.
- What should be the time-out interval ?
- The most important question about the re-transmission timer is that how long should the time-out interval be ?
- The answer to this question is difficult in the transport layer as compared to that in the data link protocol.
- Fig. 7.20.1 shows the probability density function for the time taken by data link and TCP segment acknowledgements.



(a) Data link layer **(b) For TCP**

(G-620) **Fig. 7.20.1 : Probability density of acknowledgement arrival times**

- Determining the Round Trip Time (RTT) to destination is not simple and even if we know it, deciding the value of time-out is difficult.
- Refer Fig. 7.20.1(b). If the value of time-out is too small (T_1 for example) then unnecessary re-transmission will take place.
- If time-out is too long say T_2, then the performance will degrade because re-transmission will be delayed for the long time whenever a packet is lost.
- The solution to this problem is to use a highly dynamic algorithm which adjusts the time-out interval constantly.
- This adjustment is based on continuous measurement of network performance.

Algorithms for timer management :

- Various algorithms used in TCP timer management are as follows :
  1. Jacobson's Algorithm
  2. Karn's Algorithm
  3. Persistence timer :
  4. Keepalive timer :

7.21 Comparison of UDP and TCP :

MU : Dec. 04, Dec. 05, Dec. 07, May 11, Dec. 19

University Questions

Q. 1 What do TCP and UDP use as transport layer addresses ? List the services offered by both to their upper layers. Bring out the major differences between the two. **(Dec. 04, 10 Marks)**

Q. 2 Differentiate between TCP and UDP. **(Dec. 05, Dec. 07, May 11, 4 Marks, Dec. 19, 5 Marks)**

Table 7.21.1 : Comparison of UDP and TCP

Characteristic / Description	UDP	TCP
General Description	Simple, high-speed, low-functionality "wrapper" that interfaces applications to the network layer and does little else.	Full-featured protocol that allows applications to send data reliably without worrying about network layer issues.

Characteristic / Description	UDP	TCP
Protocol Connection Setup	Connectionless; data is sent without setup.	Connection-oriented; connection must be established prior to transmission.
Data Interface To Application	Message-based; data is sent in discrete packages by the application.	Stream-based; data is sent by the application with no particular structure.
Reliability and Acknowledgments	Unreliable, best-effort delivery without acknowledgments	Reliable delivery of messages; all data is acknowledged.
Retransmissions	Not performed. Application must detect lost data and retransmit if needed.	Delivery of all data is managed, and lost data is retransmitted automatically.
Features Provided to Manage Flow of Data	None	Flow control using sliding windows; window size adjustment heuristics; congestion avoidance algorithms.
Overhead	Very low	Low, but higher than UDP
Transmission Speed	Very high	High, but not as high as UDP
Data Quantity Suitability	Small to moderate amounts of data (up to a few hundred bytes)	Small to very large amounts of data (up to gigabytes)
Types of Applications That Use The Protocol	Applications where data delivery speed matters more than completeness, where small amounts of data are sent; or where multicast/broadcast are used.	Most protocols and applications sending data that must be received reliably, including most file and message transfer protocols.
Well-Known Applications and Protocols	Multimedia applications, DNS, BOOTP, DHCP, TFTP, SNMP, RIP, NFS (early versions).	FTP, Telnet, SMTP, DNS, HTTP, POP, NNTP, IMAP, BGP, IRC, NFS (later versions).
Error control	Only checksum.	Provided.

7.22 Socket Programming :

- Many network applications consist of two programs namely a client program and a server program.
- When these programs are executed a client and a server process are created which communicate with each other by reading from and writing through the sockets.
- When creating a network application, a developer has to write the code for both client and sever programs.
- There are two different types of network applications.
- The first type of network application is an implementation of a protocol standard defined in, for example RFC.
- For such an implementation, the client and server programs must be written as per the rules of RFC.
- It is possible for two independent developers to write the client and server programs that can operate with each other properly.
- The other type of network application is a proprietary application.
- In this case the application layer protocol used by the client and server programs may not conform to any existing RFC.
- A single developer or developing team writes the client and server programs.
- As the code does not implement a public domain protocol, the other independent developers can not develop code that interoperates with the application.
- So when developing a proprietary application, the developer should not use one of the well known port numbers defined in the RFCs.

Key Issues in developing proprietary application :

- When developing a proprietary type application, the developer needs to first decide whether the application is to run over TCP or UDP.
- TCP is connection oriented and provides a reliable byte-stream channel for the data to flow between the end systems.
- The UDP is connectionless and sends data in packets between the end systems.
- But it is an unreliable protocol.
- These TCP and UDP applications are written in Java.
- It is possible to write the code in C or C++ but Java has many advantages.

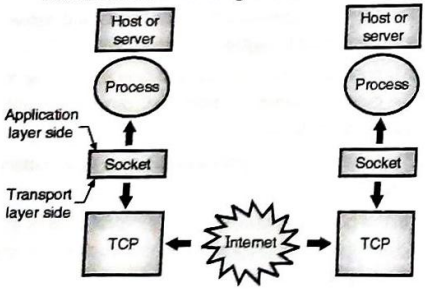
7.22.1 Socket Programming with TCP :

MU : May 16, Dec. 17

University Questions

Q. 1 Write a program for client-server application using socket programming (TCP). **(May 16, Dec. 17, 10 Marks)**

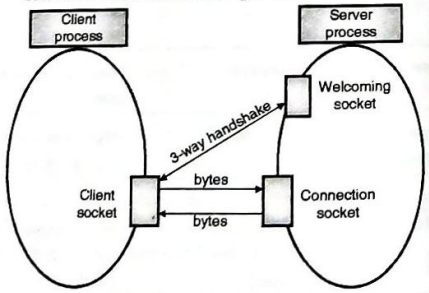
- The processes running on different machines communicate with each other by sending messages into sockets.
- This is demonstrated in Fig. 7.22.1.



- Processes are controlled by application developers
- TCP is controlled by the operating system
- UDP can be used in place to TCP

(G-630) Fig. 7.22.1 : Communicate between processes through TCP sockets

- Socket acts as a door between the application process and TCP as shown in Fig. 7.22.1.
- The application developer controls everything on the application layer side of the socket but does not have any control over the transport layer side of the socket.
- The interaction of the client and server takes place as follows.
- The client has to initiate contact with the server and when such a contact is being initiated, the server should be ready.
- That means the server must be a running process (not dormant) when a client initiates contact and the server process must have a socket to welcome the initial contact from the client.
- With the server process running, the client process can initiate a TCP connection to the server.
- This is done in the client program by creating a socket.
- When the client socket is created, the client specifies the address of the server process i.e. the IP address of the server process i.e. the IP address of the server host and the port number of the server process.
- Then the TCP on the client side initiates a three way handshake and establishes a connection with the server.
- The three way handshake and the TCP connection establishment is shown in Fig. 7.22.2.



(G-1247) Fig. 7.22.2 : Different types of sockets

- During the three way handshake the client process knocks on the welcoming socket of the server process.
- The server process responds to this knocking by creating a new socket called **connection socket** which is dedicated to that particular client.

- In the last phase of the three way handshake a TCP connection is established between the client socket and the connection socket as shown in Fig. 7.22.2.
- The TCP connection is equivalent to a direct virtual pipe between the clients socket and server's connection socket to allow a reliable byte-stream service between the client process and server process.

7.22.2 Socket Programming with UDP :

MU : Dec. 16, Dec. 18

University Questions

Q. 1 Write a program for client-server application using Socket Programming (UDP). **(Dec. 16, Dec. 18, 10 Marks)**

- As discussed in the previous section, when two processes communicate over a TCP connection, it is equivalent to communicating over a virtual pipe between the two processes.
- This pipe will remain in place until one of the processes terminates the TCP connection.
- The sending process does not have to insert the destination address to the bytes to be sent because the virtual connection is existing.
- Also the pipe provides a reliable byte transfer without altering the sequence in which the bytes are received.
- Like TCP, the UDP also allows two or more processes running on different hosts to communicate. But there is a major difference.
- The first difference is that UDP provides a connectionless service so there is no handshaking process in order to establish the virtual pipe like TCP.
- As there is no virtual pipe existing, when a process wants to send a batch of bytes to the other process, the sending process has to attach the address of the destination process.
- The destination address is a tuple consisting of the IP address of the destination host and the port number of the destination process.
- The IP address and port number together are called as **"packet"**.
- UDP provides an unreliable message oriented service in which there is no guarantee that the bytes sent by the sending process will reach the destination process.
- After creating a "packet", the sending process will push the packet into the network through a socket.
- This packet is then driven in the direction of destination process.
- The code for UDP socket programming is different than that for TCP in the following ways :
  1. No need for a welcoming socket as no handshaking is needed.
  2. No streams are attached to the socket.
  3. The sending host has to create packets.
  4. The receiving process has to obtain information from each received packet.

Review Questions

Q. 1 What do you mean by congestion control and QoS ?

Q. 2 What are the parameters of QoS ?

Q. 3 Define the term : Socket.

Q. 4 List the types of socket.

Q. 5 What are the steps used for socket programming ?

Q. 6 What are the elements of transport layer ?

Q. 7 What is difference between IP addresses and port number ?

Q. 8 What are the functions of client and server ?

Q. 9 What problems will occur in establishing a connection ?

Q. 10 What is TCP and UDP ?

Q. 11 Define threshold condition in congestion.

Q. 12 Explain the significance of listen call. Does it apply to all sockets ?

Q. 13 What parameters are specified by its various arguments.

Q. 14 Explain in detail how TCP provides flow control.

Q. 15 Define a term silly window syndrome and possible solution to overcome its effect.

Q. 16 What are the techniques used to improve QoS ?

Q. 17 What are the duties of transport layer ? Explain in brief.

Q. 18 Draw and explain the relation between network layer, transport layer and application layer.

Q. 19 What are the transport service primitives ?

Q. 20 Draw and explain the various fields of socket structure.

Q. 21 Explain connection oriented concurrent server.

Q. 22 Write a note on : Addressing in transport layer.

Q. 23 Write note on : Flow control and buffering.

Q. 24 Explain multiplexing and demultiplexing used in transport layer.

Q. 25 Explain the following issues of transport protocol : Addressing.

Q. 26 Write short notes on two-army problem in releasing a transport connection.

Q. 27 Explain Tom-Winson's three way handshake protocol to establish the transport level connection.

Q. 28 Explain how you will choose between TCP and UDP ? Compare them.

Q. 29 How does TCP tackle congestion problem using the internet congestion control algorithm.

Q. 30 Explain how TCP connections are established using the three way handshake. What happens when two hosts simultaneously try to establish a connection.

Q. 31 Explain a congestion control algorithm.

Q. 32 Explain the TCP transmission policy, congestion control.

Q. 33 Explain the following issues of transport protocol :

1. Establishing a connection.
2. Releasing a connection.

Q. 34 Give the structure of UDP header.

Q. 35 Explain the TCP header and working of the TCP protocol.

Q. 36 Explain the various fields of TCP header with the help of neat diagram.

Q. 37 Explain the various steps that are followed in releasing a TCP connection.

□□□

Module 6

Chapter 8

Application Layer

Syllabus

DNS: Name Space, Resource Record and Types of Name Server. HTTP, SMTP, Telnet, FTP, DHCP.

Chapter Contents

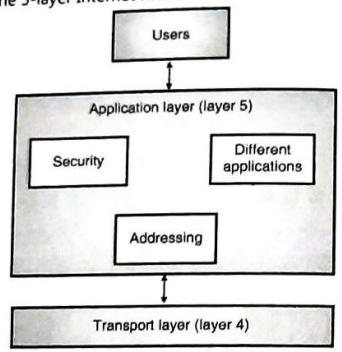
application layer.

8.1 Introduction :

- Application layer is the topmost layer in the TCP/IP protocol suite.
- The hardware and software of the Internet was designed and developed for providing various types of services at the application layer.
- All the other layers (4 of them) make these services possible.
- We will discuss various services provided at the application layer first (in this chapter) and later on study the supporting role of the other layers, providing these services.
- Many application programs have been created and used during the lifetime of the Internet. Some of them could never become standards.
- Some others have become obsolete. Some have been modified, other have been replaced by new ones.
- But some applications have survived the test of time and have become standard applications.
- Everyday new application protocols are being added to Internet.
- The Internet can provide services via two types of applications :
  1. The traditional applications.
  2. The new applications.
- The traditional applications make use of the **client server** paradigm whereas the new applications are based on the **peer-to-peer** paradigm.
- The application layer provides communication with the help of a **logical connection** which is an **imaginary connection** between the application layers of the two communicating computers. This is not the physical connection.
- The actual communication however involves all the lower layer and different types of devices such as routers, switches etc.

8.1.1 Position of Application Layer :

- The application layer is the topmost (fifth layer) of the Internet model. This is layer where all the interesting applications are found.
- People can use the Internet due to the presence of application layer.
- The layers below the application layer provide reliable transport but they do not do any real work for the users.
- In other words, the other four layers are created so that people can use the various application programs.
- Fig. 8.1.1 shows the position of application layer in the 5-layer Internet model.



(G-629)**Fig. 8.1.1 : Position of application layer**

- The application layer provides services to the users.
- The users can be humans or software. It enables the user to access the network.
- The application layer receives services from the transport layer.

Support protocols :

- For the real applications in the application layer to function, there is a need of support protocols.
- The three areas or protocols required for such support are :
  1. Network security.
  2. Domain Name Service (DNS).
  3. Network management.
- Security is not a single protocol but it contains a large number of concepts and protocols used for providing privacy.
- DNS is used to handle naming or addressing within the Internet.

- The third support protocol is network management.
- In this chapter we are going to discuss some common client - server applications that are used in the Internet.
- Some of the important applications discussed in this chapter are : DNS, FTP, HTTP, SMTP and MIME.

8.2 Providing Services :

- All the communication networks which were designed to be used in the era prior to the Internet era were designed to provide a specific type of service.
- An example of such a service is the telephone service.
- The network for telephony was originally designed to provide only the voice service.
- Later on the same network was used to provide some other services such as the FAX.
- In a similar manner, the Internet also was designed for providing service to the users all over the world.
- But the Internet is more flexible than the other services such as postal service or telephone service, due to the layered architecture of TCP/IP suite.
- Application layer being the topmost layer in the TCP/IP suite, is slightly different from the other layers.
- The application layer protocols only take services from the other layer protocols but they do not provide any service to the protocols belonging to the other layers in TCP/IP suite.
- Therefore it is easily possible to add or remove protocols to/from the application layer.
- This layer is the only layer which can provide services to the Internet users.
- Due to the flexibility of the application layer, it is possible for us to add new application protocols to the Internet.

8.2.1 Standard and Non-standard Protocols :

- The protocols belonging to the first four layers of the TCP/IP suite have to be standardized and documented in order to ensure proper operation of the Internet.
- These protocols are generally included in the package along with an operating system such as windows or UNIX.
- However the application programs can be either standard or nonstandard, for ensuring flexibility.

1. Standard Protocols (Application Layer) :

- In our day to day life, we use several application layer programs for our interaction with the Internet. These programs are standardized and well documented by the Internet authorities.
- Each standard protocol is in the form of a pair of computer programs.
- These programs have been designed to interact with the user and the transport layer so as to provide a specific service to the user.

2. Nonstandard Protocols (Application Layer) :

- By writing two programs which can interact with a user and the transport layer to provide a specific service to the user, any programmer can create a nonstandard application layer program.
- The creation of a nonstandard protocol does not need any approval of the Internet authorities if it is used privately.
- The Internet has become so popular because of these nonstandard application layer protocols.

8.3 Application Layer Paradigms :

- During the life time of the Internet, three different paradigms have been developed.
- They are as follows :
  1. Client-server paradigm.
  2. Peer to peer paradigm.
  3. Mixed Paradigm.

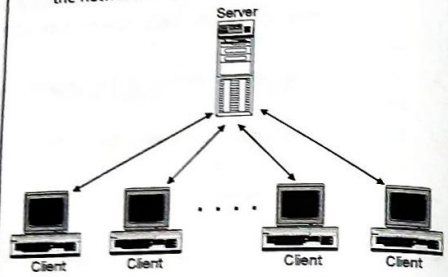
8.3.1 Traditional Paradigm : Client Server :

- The **client-server paradigm** is a traditional application layer paradigm which was the most popular paradigm until a few years ago.

Server and Client processes :

- An application program called as the server process is basically the service provider in this paradigm.

- The server process runs continuously and waits for another application program called as the **client process** to make a connection through the Internet to ask for a service.
- Some server processes have been designed to provide some specific type of services.
- The server processes are supposed to run continuously but the client process does not run continuously.
- In fact it is started when the client needs some service from a server process.
- A server process can provide the same specific service to a number of client processes which request for that service.
- In computer networking the computers connected to the Internet are known as the **end systems.**
- The examples of end systems are as follows :
  1. Desktop computers
  2. PCs
  3. Workstations
  4. Household applications
  5. Web TVs and set top boxes
  6. Digital cameras etc.
- The end systems are also known as **hosts** because they run application programs such as Web browser program, or a Web server program etc.
- Hosts can be of two different categories as follows :
  1. Client 2. Server
- In client-server network relationships, some computers act as server and other act as clients.
- A **server** is a computer, which makes the network resources available to other computers when they request it.
- It also provides some services to them. A **client** is the computer running a program that requests the service from a server.
- Local Area Networking (LAN) uses the client-server network relationship for its operation.
- You can construct a client server network by using one or more powerful computers as a servers and the remaining computers as clients.
- Client-server network typically uses a directory service to store information about the network and its users.
- All available network resources such as files, directories, applications and shared devices, are centrally managed and hosted by the server and then are accessed by client in a client-server network.
- Fig. 8.3.1 shows client-server network relationship. The server provides security and administration of the network.



(G-41) **Fig. 8.3.1 : Client server network relationship**

- In client-server networks the processing tasks are divided between clients and servers.
- Clients request services such as file storage and printing and servers deliver them.

Applications :

- The following traditional services are still using the client server paradigm for their operation :
  1. WWW : World wide web
  2. HTTP : Hyper Text Transfer Protocol
  3. FTP : File Transfer Protocol
  4. email
- Some of these protocols have been discussed in this chapter.

8.3.2 New Paradigm : Peer-to-Peer (P2P) :

- In response to the needs of some new applications on the Internet, a new paradigm called as peer to peer paradigm has emerged in recent days.
- It is also known as the **P2P** paradigm. Here the continuously running server process is not needed. Instead the responsibility of the server process is shared by the **peers.**

- Most of the Internet applications available today operate on the client-server paradigm. But gradually the **peer-to-peer (P2P) paradigm** also has gained some importance.
- The principle of P2P paradigm is that two peers (laptops, desktops or mainframes) can exchange services by communicating directly with each other.
- If the file requested by a client to server is a large file such as a music or video file, then it puts a lot of load on the server machine.
- In such situations the **P2P** paradigm becomes attractive. The P2P paradigm is also attractive in a situation in which two peers want to exchange files without involving the server.
- However it should be noted that the **P2P** paradigm does not ignore the client-server paradigm completely. Instead the P2P allows some users to share the duty of the server.
- Instead of sharing of a big file using client-server connection, the P2P paradigm will let the server download a part of that file and then share it among themselves.
- Thus in P2P paradigm the same computer has to sometimes behave like a client and at some other time like a server.
- In other words, the same computer will be a client for some applications for certain amount of time and server at other times.
- However such applications are not a part of the Internet, but they are controlled commercially.
- In P2P paradigm any computer connected to the Internet can provide service as well as request for a service.
- That means it can work as a **server** at one time and as a **client** at some other time.
- One of the best examples of Internet application in which the P2P paradigm is used is **Internet.**

Telephony :

- Another situation in which the P2P paradigm can be more useful is when one Internet users wants to share something (a file for example) with another Internet user.

Advantages :

- The main advantages of P2P paradigm are as follows :
  1. It is easily scalable.
  2. It is cost effective because an expensive server need not be used.

Disadvantages :

- Along with the above stated advantages, there are some drawbacks of P2P paradigm :
  1. Providing a secured communication is difficult.
  2. This paradigm cannot be used by all the Internet applications.

Applications :

- The following Internet applications use the P2P paradigm :
  1. Skype
  2. Internet telephony
  3. IPTV.

8.3.3 Mixed Paradigm :

- In order to get the benefits of both the paradigms, some applications may try to use the mixture of the two paradigms.

8.4 Client Server Paradigm :

- A client and a server are two running application programs called processes, and in the client server paradigm, the communication takes place at the application layer between these processes.
- A client sends a request to initialize the communication with server which is waiting for the request.
- In response to such request the server prepares a result and sends the results back to the client.
- In order to achieve this, the server should be running continuously but the client process does not have to run continuously.
- Therefore if we have two computers connected somehow to each other then we can run the client programs on one of them and the server program on the other.

Application Programming Interface.

- The server program should start running before the client program and it should run continuously. In other words, a **server** has an **infinite lifetime.**
- On the other hand a client has a **finite lifetime.**

8.4.1 Application Programming Interface (API) :

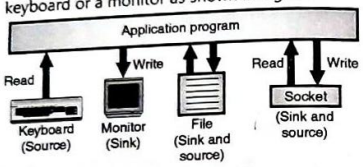
- We need to now understand that how a client process communicates with a server process.
- A client process is basically a computer program which like any other program, written in a computer language but with certain additional instructions.
- These new or additional set of instructions makes the client process capable of communicating with the other process.
- These new instructions tell the lowest four layers of the TCP/IP suite to perform the following functions :
  1. Open a connection.
  2. Send and receive data to / from the other end.
  3. Close the connection.
- A set of instructions of this type are collectively called as **Application Programming Interface (API).**
- We can define an **interface** as the set of instructions required to facilitate communication between any two entities.
- In this case one entity is the application layer process (client or server) and the other entity is the operating system.
- Therefore an API should be included while building the first four layer of an operating system, by the computer manufacturer.
- This will enable all the application layer processes to communicate with the operating system to send and receive their messages.

8.4.2 Types of APIs :

- Even though many APIs have been designed for communication, the following three are the most commonly used APIs :
  1. Socket interface
  2. Transport Layer Interface (TLI)
  3. STREAM
- Out of these the socket interface is the most common one which we will be discussing in this section.

Why Sockets ?

- With the **socket interface**, it is possible to use all the instructions that can read and write data in the programming languages like C, C++ and Java.
- All these preexisting instructions can be used to read and write data from the **sockets**.
- Thus **socket** is new source or sink of data which is being added to the programming language. It is very similar to the other sources and sinks such as a file, a keyboard or a monitor as shown in Fig. 8.4.1.



(G-2209)**Fig. 8.4.1 : Sockets are like any other sinks and sources**

8.5 Domain Name System (DNS) :

MU : Dec. 14, May 15, Dec. 17, May 19, Dec. 19

University Questions

Q. 1 Explain the need for DNS and describe the protocol functioning. **(Dec. 14, May 15, 10 Marks)**

Q. 2 Write a short notes on : DNS **(Dec. 17, May 19, 5 Marks)**

Q. 3 Explain the need for DNS (Domain Name System) and describe the protocol functioning. **(Dec. 19, 10 Marks)**

Addressing :

- For communication to take place successfully, the sender and receiver both should have addresses and they should be known to each other.
- The addressing in application program is different from that in the other layers.
- Each program will have its own address format. For example an e-mail address is like abc@vsnl.net whereas the address to access a web page is like http :// www.google.com/
- It is important to note that there is an alias name for the address of remote host.

- The application program uses an alias name instead of an IP address.
- This type of address is very convenient for the human beings to remember and use. But it is not suitable for the IP protocol.
- So the alias address has to be mapped to the IP address. For this an application program needs service of another entity.
- This entity is an application program called DNS. Note that DNS is not used directly by the user. It is used by another application programs for carrying out the mapping.

8.5.1 How does DNS Work ? : MU : May 19

University Questions

Q. 1 Write a short notes on : DNS **(May 19, 5 Marks)**

- To map a name onto an IP address, an application program calls a library procedure called the **resolver.**
- The name is passed on to the resolver as a parameter.
- The resolver sends a UDP packet to a local DNS server which looks up the name and returns the corresponding IP address to the resolver.
- The resolver then sends this address to the caller. Then the program can establish a TCP connection with the destination or sends in the UDP packets.

8.5.2 Name Space :

- The names assigned to machines should be selected carefully from the name space.
- There should be a complete control over the relation between the names and the IP addresses.
- The names and corresponding addresses are uniquely defined.
- A name space maps each address to a unique name. It can be arranged in two different ways :
  1. Flat name space. 2. Hierarchical name space.

8.5.3 Flat Name Space :

- In a flat name space, a name is assigned to every address.
- This type of name is simply the sequence of characters.
- That means it does not have any structure.
- The flat name space is not suitable for large systems like Internet, because there can be ambiguity and / or duplication.

8.5.4 Hierarchical Name Space :

- In the hierarchical name space, each name is made of many parts.
- The first part may correspond to the name of an institution, the second part may define the department and so on.
- The part that defines the nature of institution and name of institution is assigned by a central authority.
- The responsibility of deciding the rest of the name can be given to that institute itself.
- That institute can add suffix or prefix to the name for defining its host or resources.

8.6 Domain Name Space :

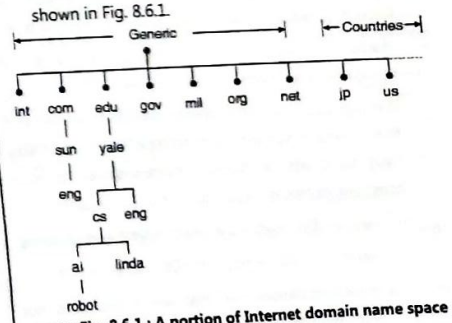
MU : Dec. 14, May 15, Dec. 19

University Questions

Q. 1 Explain the need for DNS and describe the protocol functioning. **(Dec. 14, May 15, 10 Marks)**

Q. 2 Explain the need for DNS (Domain Name System) and describe the protocol functioning. **(Dec. 19, 10 Marks)**

- Conceptually the Internet has been divided into hundreds of top level domains. Each domain covers many hosts.
- Each domain is divided into several subdomains and they are further partitioned and so on.
- These domains can be represented by a tree as shown in Fig. 8.6.1.

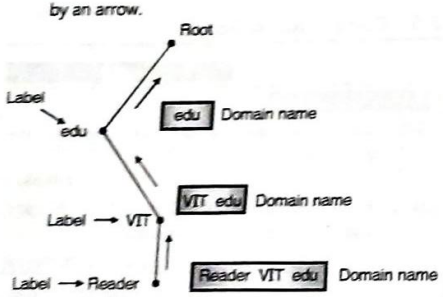


(G-631)**Fig. 8.6.1 : A portion of Internet domain name space**

- The top level domains are of two types namely generic and countries.

Generic domains :

- The generic domains are com (commercial), edu (educational institutions), gov (government), int (some international organizations), mil (military), net (network providers) and org (nonprofit organizations).
- The country domains include one entry for every country.
- Each domain is named by following an upward path. The components are separated by dots e.g. **eng.sun.com.** This is called hierarchical naming.
- Another example of hierarchical naming is shown in Fig. 8.6.2. The upward followed path has been shown by an arrow.



(G-632)Fig. 8.6.2 : Domain names, labels and hierarchical naming

Label :

- Each node in the tree has a label (or component) and it can be specified using upto 63 characters.
- If we had to remember the IP addresses of all of the Web sites we visit every day, we would all go nuts.
- Human beings just are not that good at remembering strings of numbers.
- We are good at remembering words, however, and that is where domain names come in. You probably have hundreds of domain names stored in your head. For example :

  www.yahoo.com - the world's best-known name

  www.mit.edu - a popular EDU name

  encarta.msn.com - a Web server that does not start with www

  www.bbc.co.uk - a name using four parts rather than three

  ftp.microsoft.com - an **FTP** server rather than a Web server

- The COM, EDU and UK portions of these domain names are called the **top-level domain** or **first-level domain.**
- There are several hundred top-level domain names, including COM, EDU, GOV, MIL, NET, ORG and INT, as well as unique **two-letter combinations for every country.**
- Within every top-level domain there is a huge list of **second-level domains.** For example, in the COM first-level domain, you have got :

  yahoo

  msn

  microsoft

  plus millions of others.

- Every name in the COM top-level domain must be **unique,** but there can be duplication across domains.
- For example, **msn.com** and **msn.org** are completely different machines.
- In the case of bbc.co.uk, it is a third-level domain. Up to **127 levels** are possible, although more than four is rare.
- The left-most word, such as **www** or **encarta**, is the **host name.** It specifies the name of a specific machine (with a specific IP address) in a domain.
- A given domain can potentially contain millions of host names as long as they are all unique within that domain.

Absolute and relative domain names :

- Domain names can be of two types : absolute or relative.
- An absolute domain name always ends with a dot (or period as it was called). For example eng. sun. com.
- But the relative domain does not end with a dot.

Are domain names case sensitive ?

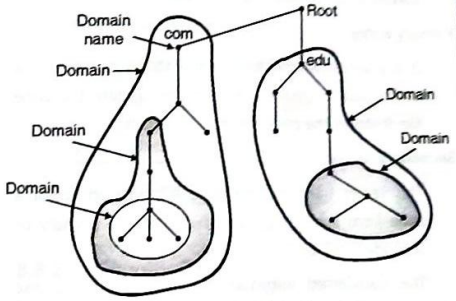
- No they are not case sensitive. So com and COM means the same thing.

How many characters ?

- Component names can have upto 63 characters and the full path name can at the most have 255 characters.
- Each domain controls how it allocates the domain under it.
- To create a new domain we have to take a permission of the domain in which it is to be included.

Domain :

- A domain can be defined as a subtree of the DNS name space as shown in Fig. 8.6.3.



(G-633)Fig. 8.6.3 : Domains

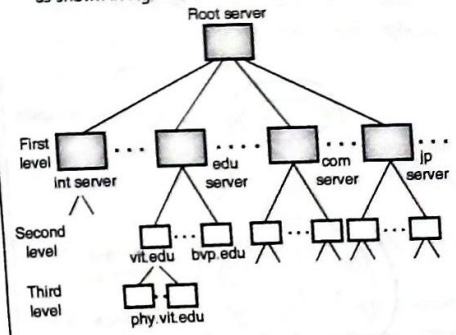
- The name of the domain is the domain name of the node at the top of the subtree as shown in Fig. 8.6.3. e.g. com or edu.
- A domain can be divided into subdomains as shown in Fig. 8.6.3.
- Note that the naming follows organizational boundaries, not physical networks.
- That means even if two different departments are located in the same building, they can have distinct domains.
- But the computers belonging to the same department kept in two different buildings will not have different domains.

8.7 Distribution of Name Space :

- The information contained in the domain name should be stored.
- But this is a huge information and if we store it on one computer then the system would be highly inefficient and unreliable.
- It will be an inefficient system because the system will be heavily loaded by the requests coming from all over the world.
- It will be unreliable because failure of one computer will make the data inaccessible. If we make a distributed name space then all these problems can be overcome.

8.7.1 Hierarchy of Name Servers :

- Name server contains the DNS database i.e. the various names and their corresponding IP addresses.
- Theoretically a single name server could contain the entire DNS database.
- But practically to store such a huge information at one place is inefficient and unreliable.
- Such a server will be soon overloaded and be useless and worst thing is if it ever goes down the entire Internet will go down.
- The solution to this problem is to distribute the information among many computers called **DNS servers.**
- Then we have to use a hierarchy of the Name servers as shown in Fig. 8.7.1.



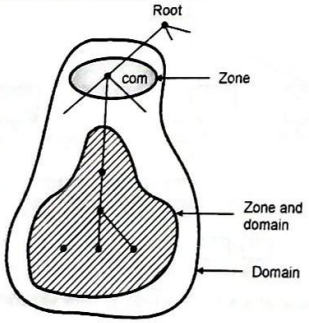
(G-634)Fig. 8.7.1 : Hierarchy of name servers

- First the whole space is divided into many first level domains.
- The root server stands alone and can create as many first level domains as required.

- The first level domains are further divided into smaller subdomains called second level domains. They can be further divided as shown in Fig. 8.7.1.
- Each server can be responsible (authoritative) to either a large or small domain.
- Note that the hierarchy of servers is similar to the hierarchy of names.
- The whole DNS name space is divided up into non overlapping **zones.** The concept of zones is as explained below.

Zones :

- With a number of DNS servers being used instead of a single one, we have to define the area over which each server has an authority.
- What a server is responsible for or has authority over is called as a zone.
- If a server is appointed for a domain and the domain is not further divided into subdomains then the domain and zone will be the same as shown in Fig. 8.7.2.
- The server makes a database called a zone file. It keeps all information about every node under that zone.
- But if a server divides its domains into subdomains and delegates a part of its authority to other servers then domain and zone will be different from each other. This is shown in Fig. 8.7.2.



(G-635)**Fig. 8.7.2 : Domains and zones**

- The information about the nodes that belong to the subdomains is stored in the servers at the lower levels.
- The higher level and original server keeps some sort of reference of these lower level servers.

Root server :

- A root server is defined as a server whose zone consists of the whole DNS tree.
- It does not store any information about domains but delegates the authority to other servers.
- It only keeps the reference of these servers.
- There are more than 13 root servers and they are distributed all around the world.

Primary and secondary servers :

- DNS defines two types of servers namely the primary servers and the secondary servers.

Primary server :

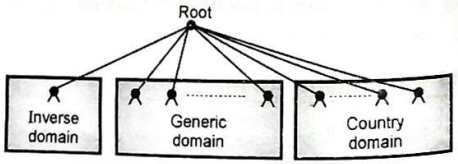
- It is a server which stores a file about its zone. It is authorised to create, maintain and update the zone file. It stores the zone file on a local disk.

Secondary server :

- This server transfers complete information about a zone from another server which may be primary or secondary server.
- The transferred information is saved on the disc storage of the secondary server.
- The secondary server is not authorized to create or update a zone file.
- If its zone file is to be updated, then it is to be done by the primary server.

8.8 DNS in the Internet :

- Let us now understand how DNS is used in Internet where the domain name space (tree) is divided into three different sections as shown in Fig. 8.8.1.
  1. Generic domain
  2. Country domain
  3. Inverse domain.



(G-636)**Fig. 8.8.1 : Use of DNS in Internet**

8.8.1 Generic Domains :

- The registered hosts are defined in the generic domains according to their generic behaviour e.g. com for commercial organizations.
- The first level in the generic domains section allows 14 possible lables. Some of them are given in Table 8.8.1.

Table 8.8.1 : Generic domain lables

Lable	Description
aero	Airline or aerospace related companies.
com	Commercial organizations.
coop	Cooperative business organizations.
edu	Educational institutions.
gov	Government institutions.
int	International organizations.
mil	Military organization.
net	Network support centers.
org	Non-profit organizations.

8.8.2 Country Domain :

- This domain section uses two character country abbreviations eg. US for united states.
- Second lable in this domain can specify organization or national designations.

8.8.3 Inverse Domain :

- The inverse domain is used for mapping an address to a name.
- This is exactly the opposite process discussed so far in which a name is mapped onto the address.

8.9 Name Address Resolution :

- The process of mapping a name to an address or vice versa is called as name address resolution.

Resolver :

- DNS application is based on the client server model. If a host wants to map a name to address or vice versa it calls a DNS client named as resolver.
- In other words, when the name ↔ address mapping is necessary a host calls a resolver.
- The resolver then sends a mapping request to the closest DNS server and accesses its storage.
- If this server has the requested information, it gives that information to the resolver but if it does not have the requested information, then it refers the resolver to other servers or asks other servers to provide the information.
- Thus the resolver receives the mapping from some source.
- It then checks for errors and if found error free delivers the mapping to the requesting process.

Mapping names to addresses :

- Generally the resolver gives a domain name to the server and requests for the corresponding IP address.
- The server checks the generic or country domains to get the corresponding address.
- If the domain name is from the generic domain section then the resolver receives a domain name such as,

  xxx.yyy.zzz.edu

- The query is sent to the local DNS server for resolution by the resolver.
- If the local server does not get the answer then, it will refer the resolver to other servers or asks them directly.
- The same procedure is followed for a name from country domain.

Mapping addresses to names :

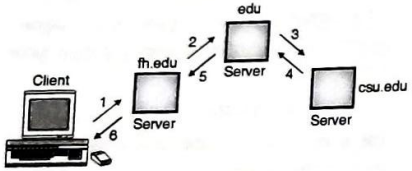
- Here, a client sends an IP address to a server and requests for its name. This type of query is called as PTR query.
- To answer the PTR query, the DNS uses the inverse domain.
- If the IP address is 142.36.48.118 then the resolver first inverts the address and adds two labels "in_addr" and "arpa" to it. So the domain name sent is :

  118.48.36.142.in_addr.arpa.

- This is received by the local DNS and resolved.

8.9.1 Recursive Resolution :

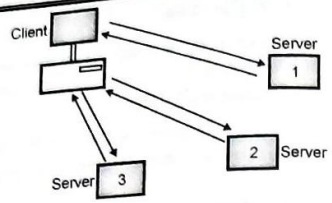
- Sometimes a client (resolver) requests for recursive or final answer from a name server.
- If this server is authorised for the domain name, it checks its database and sends a reply.
- But if this server is not authorised it diverts this request to another server (usually the parent server) and waits for the response.
- If the parent has the authority, then it sends the answer, otherwise it diverts the query to another server.
- When the query is solved, the response is returned back to the requesting client.
- Such a query is called as recursive query and the process is called recursive resolution. It is illustrated in Fig. 8.9.1.



(G-637)**Fig. 8.9.1 : Recursive resolution**

8.9.2 Iterative Resolution :

- This type of mapping can be done if the client does not ask for recursive answer.
- In iterative resolution, if the server has authority for the name it will send the answer.
- But if it does not have the authority then it returns to the client the IP address of the server that holds the answer to the query.
- The client has to repeat the query to this new server. If this server also cannot answer the query then it sends the IP address of another server to the client.
- Now the client should send the query to this third server.
- This process is called as iterative resolution because client sends the same query to different servers.
- Fig. 8.9.2 illustrates the iterative resolution.



(G-638)**Fig. 8.9.2 : Iterative resolution**

DNS examples :

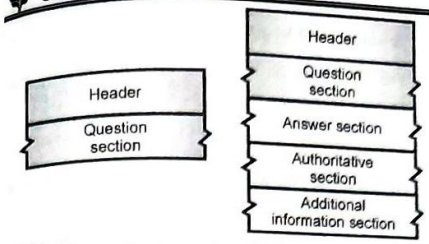
- The DNS system is a **database**, and no other database on the planet gets this many requests.
- No other database on the planet has millions of people changing it every day, either. That is what makes the DNS system so unique!

For example :

- www.yahoo.com - the world's best-known name
- www.mit.edu - a popular EDU name
- encarta.msn.com - a Web server that does not start with www
- www.bbc.co.uk - a name using four parts rather than three
- ftp.microsoft.com - an FTP server rather than a Web server
- www.spce.ac.in - Server in India 'in' domain.
- The COM, EDU and UK portions of these domain names are called the **top-level domain** or **first-level domain.**
- There are several hundred top-level domain names, including COM, EDU, GOV, MIL, NET, ORG and INT, as well as unique two-letter combinations for every country.

8.9.3 The DNS Message Format :

- DNS has two types of messages as follows and both of them have the same format.
  1. Query
  2. Responses or reply
- The formats of the two DNS messages are as shown in Fig. 8.9.3.



(a) Query **(b) Response or reply**

(G-639)**Fig. 8.9.3**

- Both query and reply messages have the same header format with some fields set to zero for query messages.
- The header is 12 byte long. The header format for both the types of messages is shown by shaded portions in Fig. 8.9.3.

8.9.4 Caching :

- Every time a query is asked, the server has to spend time in searching the corresponding IP address.
- If this searching time is reduced then efficiency would go up.
- The searching time can be reduced by using a technique called caching.
- When a server asks for a mapping from another server and receives the response, it stores this information in its cache memory before sending it to the client.
- If the same or other client request for the same mapping, it can check its cache memory and resolve the problem at its own level.
- This will certainly save a lot of time.
- But the problem with caching is that, if a server caches (stores) a mapping for a long time then the mapping may get outdated and the client will not get the latest mapping.
- This problem can be solved by adding the time to live information (TTL) to the mapping and each server is asked to keep a TTL counter for each mapping in its cache.

8.9.5 DNS Records : MU : Dec. 13

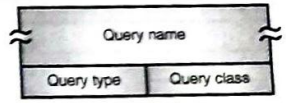
University Questions

Q. 1 What is a resource record ? How it is useful for DNS ? **(Dec. 13, 7 Marks)**

- There are two types of records used in DNS as follows :
  1. Question records and
  2. Resource records
- The question records are used in the query and response messages whereas the resource records are used in the response messages.

Question Records :

- The client uses the question record to get the required information from the server.
- The format of question record has been shown in Fig. 8.9.4.



(G-1791) **Fig. 8.9.4 : Question record format**

- Various fields in the question record format are query types and query class.

Query name :

- This field has a variable length and it contains a domain name.
- The count field tells us how many characters are present in each section.

Query type :

- This field is 16-bit long and it defines the type of query.
- Some of the commonly used query types are A, NS, CNAME, SOA, ANY etc.

Query class :

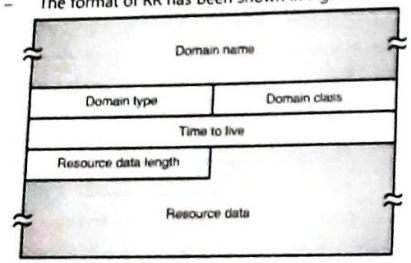
- This field also is 16-bit long. It defines the specific protocol using DNS.
- Table 8.9.1 has listed some of possible classes. However the most important class would be IN i.e. internet (class 1).

Table 8.9.1 : Query classes

Class	Mnemonic	Explanation
1.	IN	Internet
2.	CSNET	CSNET network (Not used now)
3.	CS	COAS network
4.	HS	The Hesiod server (MIT)

Resource Record :

- Each domain name i.e each node on the tree in DNS is associated with the resource record which is a part of the server database.
- Resource records are returned by the server to the client.
- The format of RR has been shown in Fig. 8.9.5.



(G-1792) **Fig. 8.9.5 : Format of resource record**

1. Domain name :

- This field contains the domain name and its length is not fixed. It has a variable length.
- The domain name in the question record is duplicated here.

2. Domain type :

- This field and the query type field in the question record are the same except the last two types i.e. AXFR and ANY are not allowed.

3. Domain class :

- This field is same as the query type field in the question record.

4. Time to - live :

- This field is 32 bit long and it defines the time for which the answer is valid (in seconds).
- If the contents of this field is zero, then it indicates that the resource record is used only in a single transaction.

5. Resource data length :

- This field is 16 bit long. It is used for defining the length of the resource data.

6. Resource data :

- As shown in Fig. 8.9.5 the resource data field is a variable length field.
- It contains the answer to the query or domain name or the additional information.
- The format and contents of this field depend on the value of the type field. It can be one of the following :
  1. A number
  2. A domain name
  3. An offset pointer
  4. A character string.

Encapsulation :

- DNS can use either TCP or UDP. It may choose any one of these protocol but in either case the server uses port 53.
- The UPD is preferred if the length of response message is upto 512 bytes whereas TCP is used if the message length is larger than 512 bytes.

Registrars :

- New domains are added to DNS through a registrar, which is a commercial entity.
- Whenever an organization applies for DNS domain name, a registrar first checks that the requested domain name is unique.
- The applying organization has to give the name of its server and IP address of the server to the registrar.

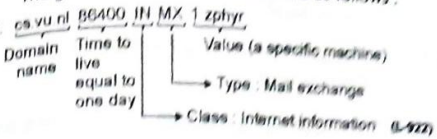
Ex. 8.9.1 : Explain Domain Resource Records. Describe following DNS Resource Records.

1. cs.vu.nl 86400 IN MX 1 zphyr
2. flits 86400 IN A 130.37.20.10

Soln. :

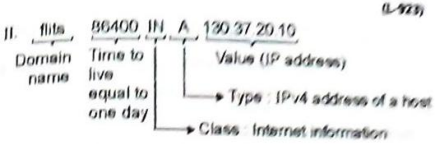
- The general format of a DNS resource record is as follows :
  1. Domain-name
  2. Time-to-live
  3. Class Type Value.
- The given DNS resource records are as follows :

I. cs.vu.nl 86400 IN MX 1 zphyr



(G-922)

II. flits 86400 IN A 130.37.20.10



(G-923)

8.10 World Wide Web (WWW) :

- People have become aware of the power of Internet through WWW.
- HTTP is a file transfer protocol which is specifically designed to facilitate access to the WWW.
- The World Wide Web is an architectural framework for accessing documents which are spread out over a number of machines over Internet.
- It has a colourful graphical interface which is easy for the beginners to use.
- It provides information on almost every subject. The web (also known as WWW) began in 1989 at CERN the European center for nuclear research.
- The web was designed basically to connect scientists stationed all over the world. The web is basically a client-server system.
- The web pages are written in the languages HTML and Java.
- The growth of the World-Wide Web (WWW or simply Web) today is simply phenomenal.
- Each day, thousands of more people join the Internet (above 100 million users at recent estimates).
- Easy retrieval of electronic information along with the multimedia capabilities of Web browsers (like Mosaic or Netscape) are the factors responsible for this explosion.
- This topic provides some basic information behind some of this technology used in accessing the World-Wide Web.

Difference between Web and Internet :

- The Web and the Internet are not the same thing.
- The Web is a collection of standard protocols or instructions, sent back and forth over the Internet to gain access to information.
- The Internet, on the other hand, is a "network of networks" -- a more physical entity.

8.11 HTTP (Hypertext Transfer Protocol) :

MU : Dec. 14

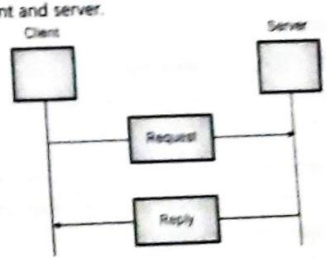
University Questions

Q. 1 Give short notes on : HTTP (Dec. 14, 5 Marks)

- The main function of HTTP is to access data on WWW.
- This protocol can access the data in various forms such as plaintext, hypertext, audio, video etc.,
- The function of HTTP is equivalent to a combination of FTP and SMTP.
- It uses services of TCP. It uses only one TCP connection (port 80).
- There is no separate control connection like the one in FTP.
- Only the data transfer takes place between the client and server so there is only one connection and it is the data connection.
- The data transfer in HTTP is similar to SMTP. The format of the messages is controlled by MIME like headers.

8.11.1 Principle of HTTP Operation :

- The principle of HTTP is simple. A client sends a request. The server sends a response.
- The request and response messages carry data in the form of a letter with a MIME like format.
- Fig. 8.11.1 shows the HTTP transactions between client and server.



(G-657) **Fig. 8.11.1 : HTTP transaction**

- The client initializes the transaction by sending a request message and the server responds by sending a response.

8.11.2 The Web and HTTP :

- HTTP is the Web's application layer protocol. It is the heart of the Web.
- It has been defined in [RFC 1945] and [RFC 2616].
- HTTP is implemented in two programs :
  1. A client's program 2. A server's program.
- These programs are executed on different and systems and talk to each other by exchanging HTTP messages.
- HTTP defines how Web clients such as browsers request Web pages from Web servers and how servers transfer Web pages to clients.
- HTTP uses TCP as its under lying transport protocol (rather than using UDP).
- The HTTP client first initiates a TCP connection with the server.
- After establishing a connection, the browser and the server processes access TCP through their socket interface.
- TCP provides a reliable data transfer service to HTTP.
- That means each HTTP request message, transmitted by a client will eventually arrive intact at the server.
- Similarly each HTTP response message transmitted by the server will eventually arrive intact at the client, due to the reliable TCP connection.
- Due to this kind of layered architecture HTTP need not have to worry about the lost data or about the details of how TCP deals with the loss and retransmission of data. It is managed by TCP.

Statelessness :

- In HTTP, the server sends the files requested to the client without storing any state information about the client.
- So it may happen that the same client may ask the same information repeatedly to the server and the server would not even understand it.
- So it will keep resending those files.
- As the HTTP servers does not maintain any information about the state of client it is called as a **stateless protocol.**

8.11.3 Non-persistent and Persistent Connection :

- HTTP is capable of using both non-persistent and persistent connections.
- HTTP uses persistent connection in its default mode.
- But HTTP clients and servers can be configured to use the non-persistent connection as well.

1. Non-persistent connections :

- Let us discuss the step-by-step procedure followed for transferring a web page from server to client for a non-persistent connection.
- Imagine that the web page consists of a base HTML file and many JPEG images and that all these objects reside on the same server.
- Let the URL for the base HTML file be as follows :

  http://www.vit.edu/itdept/home.index
- Then the sequence of events is as follows :
  1. The HTTP client process initiates a TCP connection to the server www.vit.edu on port number 80, which is the default port number for HTTP.
  2. The HTTP client, sends an HTTP request message to the server via its socket associated with the TCP connection. This request message is of the following format :

     Path name/itdept/home.index.
  3. The HTTP server process receives the request message via its socket associated with the connection. It then retrieves the object.

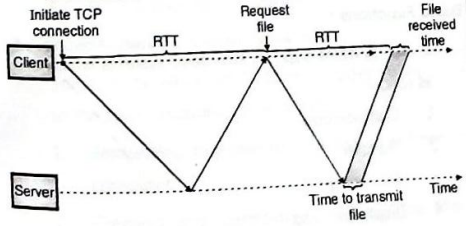
     /itdept/home.index
- from its storage. It then encapsulates this retrieved object in an HTTP response message and sends the response message to the client via its socket.
  4. The HTTP server process tells TCP to close the TCP connection.
  5. As soon as the HTTP client receives the response message, the TCP connection is terminated.

6. The response message indicates that the encapsulated object is an HTML file. The client takes out the file from the response message and examines the HTML file. The client will find references to all the JPEG objects.
7. The client follows the first four steps for each JPEG object.

- As the browser receives the web page, it displays the page.
- Different browsers can display the same web page differently.
- However HTTP is not concerned about this. Its specifications define only the communication between the HTTP client program and HTTP server program.
- The steps discussed earlier were for the **non-persistent** connection where each TCP connection is closed after the server sends the object.
- That means the TCP connection does not persist for other objects.
- Each TCP connection transports one request message and one response message.

Round-Trip Time (RTT) :

- The RTT is defined as the time taken by a small packet to travel from client to server and then back to the client.
- The components of RTT are :
  1. Packet propagation delays
  2. Packet queuing delays 3. Packet processing delays.
- Now consider the sequence of events taking place when a user clicks on a hyperlink.
- These events are illustrated in Fig. 8.11.2.



(G-658) **Fig. 8.11.2**

1. The browser initiates a TCP connection between the browser and web server. This process makes use of a **three way handshake**.
- In the three way handshake, the client sends a small TCP segment to the web server.
- The server acknowledges and responds with another small TCP segment. Finally the client acknowledges back to the server.
2. After completing the first two parts of the three way handshake the client sends the HTTP request message to the server.
3. In response the server sends the HTML file to the client. The total response time as shown in Fig. 8.11.2 is equal to 2RTT plus the time taken by the server to transmit the file.

Disadvantages of non-persistent connections :

1. It is necessary to establish and maintain a new connection for each requested object.
2. For each connection TCP buffers need to be allocated and TCP variables need to be kept in both the client and server.
3. There is a delay of 2RTTs associated with the transfer of each object.

2. Persistent connection :

- The disadvantages of non-persistent connections can be overcome if persistent connection is used.
- With the persistent connection, the server leaves the TCP connection open after sending a response.
- All the requests and responses between the same client and server can be sent over the same connection.
- Hence the entire web page can be sent over a single persistent connection.
- It is also possible to send the multiple web pages reciding on the same server to the same client over a single persistent TCP connection.
- The TCP connection is closed only after the time out interval by the HTTP server.

Types of persistent connections :

- The two versions of persistent connections are as follows :
  1. Without pipelining 2. With pipelining.

1. Without pipelining :

- For this version, the client has to issue a new request only when it receives the previous response.
- The delay of only one RTT is experienced by the client in order to request and receive each object.
- This is an improvement over the non-persistent connection which experiences a delay of 2RTT. This delay can be reduced by using pipelining.
- Another disadvantage of no pipelining is that the TCP connection becomes idle i.e. does nothing while it waits for another request after the server had sent an object.

2. With pipelining :

- This mode reduces the delay further. The default mode of HTTP uses persistent connection.
- With pipelining the HTTP client will issue a request as soon as it encounters a reference.
- This allows the HTTP to make back to back requests. It can make a new request before receiving the response.
- When the server receives back to back requests, it sends the objects back to back.
- With pipelining only one RTT will be expended for all the referenced objects.
- Another advantage is that the pipelined TCP connection remains idle for a very short time.

8.11.4 HTTP Messages :

- The HTTP messages are of two types :
  1. Request message 2. Response message.
- The format of both these messages is almost the same.

8.12 Electronic Mail :

- One of the most popular network services is electronic mail (e-mail).
- Simple Mail Transfer Protocol (SMTP) is the standard mechanism for electronic mail in the internet.
- The first e-mail systems simply consisted by file transfer protocols.
- But some of the limitations of this system were as follows :
  1. It is difficult to send a message to a group of people.
  2. Message did not have any internal structure. So its computer processing was difficult.
  3. The sender never used to know if a message arrived or not.
  4. It was not easy to handover one's e-mails to someone else for the purpose of managing them when one is out of town or country for sometime.
  5. The user interface with the transmission system is poorly integrated.
  6. It was not possible to create and send messages containing a text, drawing, facsimile and voice together.
- So more elaborate e-mail systems were proposed. ARPANET e-mail proposals were published as RFC 821 (transmission protocol) and RFC 822 (message format).
- These are used in Internet.

8.12.1 E-mail Architecture and Services :

- An e-mail system consists of two subsystems :
  1. User agents and
  2. Message transfer agents.

User agents : They enable users to read and send e-mail.

Message transfer agents : They move the messages from the sender to the receiver.

Basic Functions :

- E-mail systems support five basic systems which are as follows :
  1. Composition
  2. Transfer
  3. Reporting
  4. Displaying and
  5. Disposition

1. Composition :

- The process of creating messages and to answer them is known as composition.
- The system can also provide assistance with addressing and a number of header fields attached to each message.

2. Transfer :

- It is the process of moving messages from the sender to the recipient.
- This includes establishment of a connection from sender to destination or some intermediate machine, transferring the message, and breaking the connection.

3. Reporting :

- The reporting system is designed to tell the sender about whether the message was delivered or rejected or lost.

4. Displaying :

- It is the process of displaying the incoming messages so that it can be read by the user.
- For this purpose simple conversions and formatting are required to be done.

5. Disposition :

- This is concerned with what the recipient does with the received message. Disposition is the final step in e-mail system.
- Some of the possibilities are as follows :
  1. Throw after reading
  2. Throw before reading
  3. Save messages
  4. Forward messages
  5. Process messages in some other way.

Advanced features of E-mail systems :

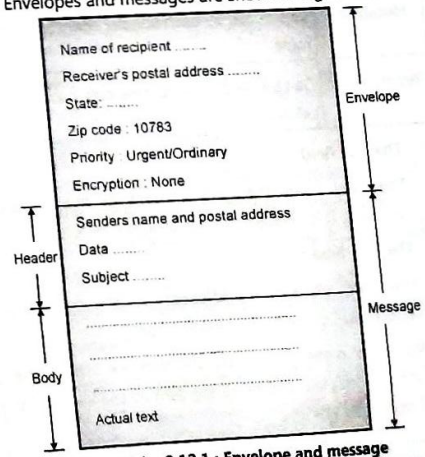
- Some of the advanced features included in addition to the basic functions are as follows :
  1. Forwarding an e-mail to a person away from his computer.
  2. Creating and destroying mailboxes to store incoming e-mail.
  3. Inspecting contents of mailbox, insert and delete messages from the mailboxes.
  4. Sending a message to a large group of people using the idea of mail list.
  5. To provide the facility of registered e-mail.
  6. Automatic notification of undelivered e-mails.
  7. Carbon copies
  8. High priority e-mail (setting the priority of e-mails)
  9. Secret (encrypted e-mail)
  10. Alternative recipient. This allows automatic forwarding of an e-mail to an alternate recipient if the main recipient is not available.

E-mail Envelope :

- In the modern e-mail systems, there is a distinction made between the e-mail and its contents.
- An e-mail envelope contains the message, destination address, priority, security level etc.
- The message transport agents such as SMTP use this envelope for routing.

Message :

- The actual message inside the envelope is made of two parts :
  1. Header and 2. Body
- Header carries the control information while body contains the message contents.
- Envelopes and messages are shown in Fig. 8.12.1.



(G-640) Fig. 8.12.1 : Envelope and message

8.12.2 Message Formats :

- Let us now discuss the e-mail message formats.

RFC 822 :

- All the e-mail messages consist of an envelope, a few header fields, a blank line and then the message body.
- Each header field logically consists of a single line of ASCII text which consists of the field name, a colon and a field.
- Normally the user agent builds a message and passes it to the message transfer agent which uses some header fields for construction of an envelope.
- Table 8.12.1 shows the principle header fields related to the message transport. Let us discuss them one by one.

Table 8.12.1 : RFC 822 header fields related to message transport

Header Name	Meaning
To :	E-mail address of primary recipients
Cc :	E-mail address of secondary recipients (Carbon copy)
Bcc :	E-mail address for blind carbon copies
From :	Originator of the message
Sender :	E-mail address of the person sending the message
Received :	Line added by each transfer agent along the route
Return – Path :	Can be used to identify the path back to the sender.

1. **The To : field :**
   - This field gives the DNS address of the primary recipient. It is allowed to have multiple recipients.
2. **The Cc : field :**
   - This field gives the addresses of any secondary recipients. Cc stands for carbon copy.
   - Whatever message and attachments are sent to the primary recipient the same are sent to the secondary recipient as well.
3. **The Bcc : field :**
   - The long form of Bcc is blind carbon copy. This field is like Cc field, except that this is deleted from all the copies sent to the primary and secondary recipients.
   - Thus a sender can send copies to third parties without primary and secondary recipients knowing about it.
4. **From : and Sender : fields :**
   - These fields tell about who wrote the message and who actually sent the message respectively because the person who creates the message and the person who sends it can be different.
   - The From : Field is necessary but the Sender : field can be omitted, if it is same as the From : field.
   - These fields are required when the message cannot be delivered and is to be returned to the sender.
5. **Received : field :**
   - A line containing Received : is added by each message transfer agent along the way.
   - This line carries the agent's identity, date and time at which the message was received.
   - It also contains some other information that can be used to find bugs in the routing system.
6. **The Return-Path : field :**
   - This field is added by the final message transfer agent and it is intended to tell how to get back to the sender.
   - This information can be obtained from all the received : headers.

Other header fields :

- In addition to the fields of Table 8.12.2, RFC 822 messages may contain many other header fields.
- These are used by either the user agents or human recipients some of them are shown in Table 8.12.2.

Table 8.12.2 : Some fields in RFC 822 message header

Header	Meaning
Date :	The date and time of the message.
Reply-To	E-mail address to which the reply is to be sent
Message-Id :	Message identifying number

Header	Meaning
In-Reply-To :	Message-Id of the message to which this is a reply
References :	Other relevant message identifying numbers
Keywords :	Keywords chosen by user
Subject :	Summary of the message for the one line display.

- The RFC 822 allows the users to invent new headers for their own private use but it is essential that these headers start with the string X-. For example X-Event of the week.

Message Body :

- The message body comes after the header. The users can include anything that they want to send, in the message body.
- It is possible to terminate the messages with ASCII cartoons, quotations, political statements etc.

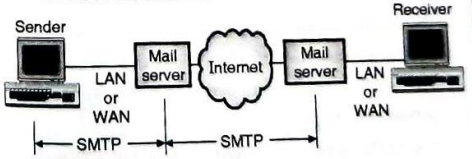
8.13 Message Transfer Agent : SMTP :

MU : Dec. 16, Dec. 17

University Questions

Q. 1 Write short notes on : SMTP. **(Dec. 16, Dec. 17, 5 Marks)**

- The actual mail transfer is carried out through the message transfer agent.
- A system should have the client MTA in order to send a mail and it should have a server MTA in order to receive one.
- SMTP is the protocol which defines MTA client and server in the Internet.

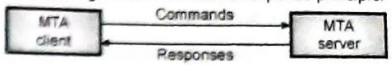


(G-641)Fig. 8.13.1 : SMTP range

- As shown in Fig. 8.13.1, the SMTP is used twice, once between the sender and sender's mail server and then between the two mail servers.
- The job of SMTP is simply to define how commands and responses be sent back and forth.
- Each network can choose its software package for implementation.

8.13.1 Commands and Responses :

- As shown in Fig. 8.13.2, SMTP the transfer of messages between MTA client and MTA server takes place using the command and response principle.



(G-642)Fig. 8.13.2 : Commands and responses in HTTP

- Each command or response is terminated by a two character end of line token.
- The two characters used are carriage return and line feed.

8.13.2 SMTP (Simple Mail Transfer Protocol) :

MU : Dec. 18, May 19

University Questions

Q. 1 Write a short note on SMTP. **(Dec. 18, 10 Marks, May 19, 5 Marks)**

- In internet the source machine establishes a connection to port 25 of the destination machine so as to deliver an e-mail.
- An e-mail daemon which speaks SMTP is listening to this port.
- This daemon is supposed to perform the following tasks :
  1. Accept the incoming connections, and copy messages from them into appropriate mailboxes.
  2. Return an error message to the sender, if a message is not delivered.
- SMTP is a simple ASCII protocol.
- Once a TCP connection between a sender and port 25 of the receiver is established, the sending machine operates as a client and the receiving machine acts as a server.
- The client then waits for the server to take initiative in communication.
- The server sends a line of text which decares its identity and announces its willingness/ unwillingness to receive mail.

- If the server is not prepared, the client will release the connection, wait for sometime and try again later.
- But if the server is willing to accept e-mail, then the client announces the sender of e-mail and its recipient.
- If such a recipient exists at the destination, then the server tells the client to send the message.
- The client, then sends the message and the server sends back its acknowledgement.
- No checksums are generally required because TCP provides a reliable byte stream. If there are any more e-mail, then they can be sent now.
- After exchanging all the e-mail, the connection is released.
- SMTP uses numerical codes. The lines sent by the client are marked C : ; and those sent by the server are marked S : ;
- Some of the commands, useful for communication are :

  HELO, RCTP, DATA, QUIT etc.
- RCTP represents recipient. If only one command is used then the message is being sent to only one recipient.
- If the command is used many times, then it indicates that the message is sent to more than one recipients.
- In such a case each message is individually acknowledged or rejected.
- The syntax of four character commands for the clients are rigidly specified but the syntax for the replies are not that rigid.
- The SMTP protocol is well defined by RFC 821 but some problems are still present.

Problems in SMTP :

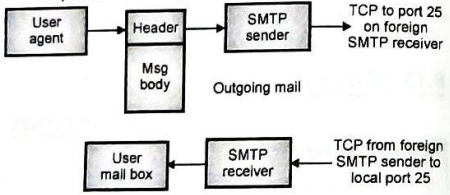
- Some of the problems in SMTP are as follows :
  1. Some older versions of SMTP are not capable of handling messages longer than 64 kB.
  2. If client and server have different time-outs, then one of them may give up when the other is still busy. This will terminate the connection unnecessarily.
  3. In rate situations, infinite mailstorms can be triggered.

Extended SMTP (ESMTP) :

- Some of these problems can be solved by using the extended SMTP (ESMTP) which is defined in RFC 1425.

8.13.3 Components of E-mail System :

- The three main components of internet mail system are :
  1. User Agent (UA)
  2. SMTP sender
  3. SMTP receiver
- They are shown in Fig. 8.13.3.
- The mail is created by a user agent program in response to user input.



(G-643) **Fig. 8.13.3 : SMTP mail flow**

- Each created message consists of a header which includes the recipient's E-mail address and other information and the message body containing the message to be sent.
- These messages are lined up to form a queue and provided as input to an SMTP sender program.
- The SMTP sender takes messages from the queue and transmits them to the proper destination host via SMTP connection over one or more TCP connections to port 25.
- The SMTP protocol is used to transfer a message from the SMTP sender to SMTP receiver and it uses TCP connection for the same.
- The SMTP receiver accepts each arriving message and stores it in the user mail box.
- If the mail is to be forwarded then the SMTP receiver copies it to the outgoing mail queue.

8.13.4 SMTP Commands :

- The operation of SMTP consists of a series of commands and responses exchanged between the SMTP sender and receiver.
- The SMTP sender establishes the TCP connection to the receiver.
- After establishing the connection, the SMTP sender sends commands over the connections to the receiver.
- The SMTP receiver generates exactly one reply from the SMTP receiver.
- Table 8.13.1 shows the SMTP commands.
- Each command consists of a single line of text which begins with a four letter command code followed in some cases by an argument field.
- Most replies are a single line. However multiline replies also are possible.

Table 8.13.1 : SMTP commands

Name	Description
HELO	Send identification of the sender.
MAIL	Identifies originator of mail.
RCPT	Identifies recipient of mail.
DATA	Transfer message text.
RSET	Abort the current mail transaction.
NOOP	No operation.
QUIT	Close TCP connection.
SEND	Send mail to terminal.
SOML	Send mail to the terminal if possible, otherwise to mailbox.
SAML	Send mail to terminal and mail box.
VRFY	Confirm user name.
EXPN	Return membership of mailing list.
HELP	Send system-specific documentation.
TURN	Reverse role of sender and receiver.

8.13.5 SMTP Operation :

- The basic SMTP operation occurs in three phases :
  1. Connection setup
  2. Exchange of one or more command-response pairs
  3. Connection termination

1. Connection setup :

- The sender opens (i.e. creates) a TCP connection with the receiver.
- Once the connection is established, the receiver identifies itself with "220 Service Ready".
- The sender identifies itself with HELO command.
- The receiver accepts the sender's identification with "250 OK".

2. Mail transfer :

- Once the connection has been established, the SMTP sender may send one or more messages to SMTP receiver.
- There are three logical phases to transfer a message :
  1. A MAIL command identifies the originator of message.
  2. One or more RCPT commands identify the recipient for this message.
  3. A DATA command transfers the message text.

3. Connection closing :

- The SMTP sender closes the connection in two steps. First the sender sends a QUIT command and waits for a reply.
- Second step is to initiate a TCP close operation for the TCP connection.
- The receiver initiates its TCP close after sending its reply to the QUIT command.

8.13.6 Comparison of HTTP and SMTP :

Table 8.13.2 : Comparison of HTTP and SMTP

Sr. No.	SMTP	HTTP
1.	Message is transferred from client to server.	Message transfer is from client to server or the other way round.
2.	Uses TCP.	Uses TCP.

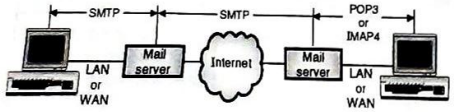
Sr. No.	SMTP	HTTP
3.	Uses port 25 for transmission.	Uses port 80 for transmission.
4.	SMTP messages are to be read by humans.	HTTP messages are to be read and understood by the HTTP servers and HTTP clients.
5.	These messages are first stored and then forwarded.	These messages are immediately delivered.

8.14 Message Access Agent : POP and IMAP :

- The SMTP is used in the first and second stages of mail delivery.
- But SMTP is not used in the third stage, because SMTP is a push protocol which is meant for pushing the message from client to server.
- The third stage needs a **pull** protocol because the client has to pull messages from the server.
- The bulk data gets transferred from the server to client.
- Therefore third stage uses a message access agent which is a pull protocol.
- The two message access agents available are :
  1. Post Office Protocol, version 3 (POP 3).
  2. Internet Mail Access Protocol (IMAP 4).

8.14.1 POP 3 :

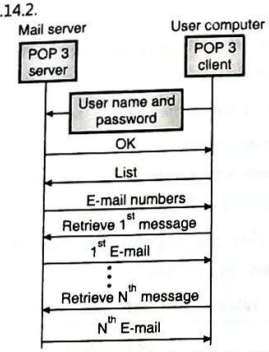
- The POP3 consists of client POP3 software and server POP3 software.
- Out of these, the client POP3 software is installed on the receiving computer whereas the mail server gets the server POP3 software installed on it.
- When the user wants to download email from the mailbox on the email server, the events take place in the following sequence. Refer Fig. 8.14.1.



(G-645) **Fig. 8.14.1 : Use of POP 3 or IMAP 4**

1. The client (user) establishes a connection with the server on TCP port 110.
2. The client then sends its user name and password to the server in order to access the mailbox.
3. The user is then allowed to list and get the mail messages one by one.

- This is called as downloading. It is illustrated in Fig. 8.14.2.



(G-647) **Fig. 8.14.2 : Downloading in POP3**

Modes of POP 3 :

- POP3 has two modes of operation :
  1. Delete mode and 2. Keep mode.
- **Delete mode :** In this mode the mail is deleted from the mailbox after each retrieval.
- This mode is used when the user is working on his permanent computer because it is then possible for him to save and rearrange the received mail after reading it.
- **Keep mode :** If operated in this mode, the mail remains in the mailbox after retrieval.
- This mode is used when the user accesses mail away from the primary computer. The read mail can be organized later.

Disadvantages of POP3 :

1. POP3 does not allow organization of email on the server.
2. The user can not create different folders on the server. It can create them only on his own computer.

3. The user can not partially check the contents of E-mail before down loading.

8.14.2 IMAP4 :

- Internet Mail Access Protocol Version 4 (IMAP4) is another mail access protocol which is very similar to POP3 but has more features.
- This makes IMAP4 more powerful but more complex as compared to POP3.
- IMAP is more sophisticated than POP3 and it is defined in RFC 1064.
- IMAP is ideal for a user having multiple computers such as a laptop on the road, PC at home and a workstation in office.
- IMAP maintains a central repository which can be accessed from any machine.
- So IMAP does not copy e-mail to the user's personal machine.
- An important feature of IMAP is its ability to address mail not by arrival number but by using attributes.
- That means the mailbox is like a relational database system than a linear sequence of messages.

Extra features of IMAP4 :

1. It is possible for the user to check the header before down loading.
2. It is possible for the user to search for the contents of E mail before downloading.
3. It is possible to partially download E mail.
4. It is possible for the user to create, rename or delete mailboxes on the mail server.
5. It is possible for the user to create a hierarchy of mailboxes in a folder for storing e-mails.

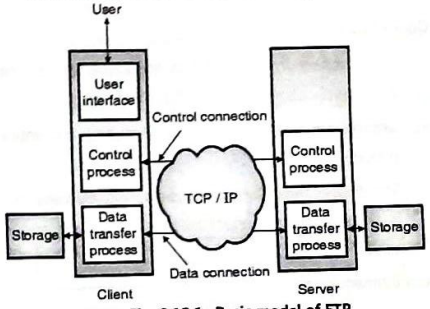
8.15 File Transfer Protocol (FTP) :

MU : Dec. 19

University Questions

Q. 1 Write short note on FTP. **(Dec. 19, 10 Marks)**

- A standard mechanism provided by the Internet which helps in copying a file from one host to the other is known as the File Transfer Program (FTP).
- Some of the problems in transferring files from one system to the other are as follows :
  1. Two systems may use different file name conventions.
  2. Two systems may represent text and data in different ways.
  3. The directory structures of the two systems may be different.
- FTP provides a simple solution to all these problems.
- The basic model of FTP is shown in Fig. 8.15.1.



(G-648) **Fig. 8.15.1 : Basic model of FTP**

- FTP establishes two types of connections between the client and server.
- One of them is used for data transfer and the other is for the control information.
- The fact that FTP separates control and data makes it very efficient.
- The control connection uses simple rules of communication.
- Only one line of command or a line of response is transferred at a time.
- But the data connection uses more complex rules due to the variety of data types being transferred.
- FTP uses port 21 for the control connection and port 20 for the data connection. Both these are well known TCP ports.
- As shown in Fig. 8.15.1 the client is made of three blocks namely :
  1. User interface
  2. Control process and
  3. Data transfer process.

- The server has two blocks : the control process and data transfer process.
- The control connection connects the control processes while data connection connects the data transfer processes as shown in Fig. 8.15.1.
- The control connection is kept alive during the entire interactive FTP session.
- The data connection is first opened, file is transferred and data connection is closed. This is done for transferring each file.

Control connection :

- This connection is created in the same way as the other application programs described earlier.
- Control connection remains alive during the entire process.
- The IP uses minimize delay type service because this is an interactive connection between a user and a server.

Data connection :

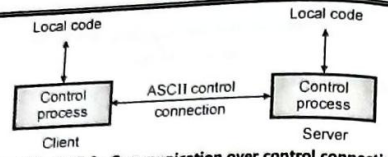
- Data connection uses the port 20 at the server site.
- This connection is opened when data to be transferred is ready and it is closed when transfer of data is over.
- The data connection does not remain open continuously like control connection. It is opened and closed many times as per requirement.

8.15.1 Communication in FTP:

- FTP operates in client – server environment.
- The two computers involved in communication may be different in terms of the operating systems, character sets, file structures and file formats etc.
- FTP can make them compatible.
- The approaches for communication over control connection and data connection are different from each other.

1. Communication over control connection :

- Refer Fig. 8.15.2 to understand the FTP's approach for the communication over the control connection.

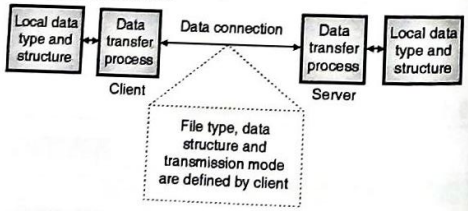


(G-649)**Fig. 8.15.2 : Communication over control connection**

- Similar to SMTP, FTP uses a set of ASCII characters to communicate across the control connection.
- Communication is achieved through a process of commands and response. One command is sent at a time.
- Each command or response is only of one short line.
- So it is not necessary to think about file format or file structure.
- Each line is ended with a two character token. The two characters used in the token are carriage return and line feed.

2. Communication over data connection :

- The purpose of implementing a data connection is to transfer a file.
- For this the client has to define the following :
  1. Type of file being transferred.
  2. Structure of data in the file
  3. Mode of transmission.
- Before the transmission over data connection, the communication over control connection is performed.
- Refer Fig. 8.15.3 to understand communication over data connection.



(G-650)**Fig. 8.15.3 : Communication over the data connection**

- The problem of heterogeneity is solved by defining three attributes of communication : file type, data structure and transmission mode.

8.15.2 File Types :

- FTP can use one of the following file types for transfer of data over the data connection :
  1. ASCII file 2. EBCDIC file 3. Image file.
- ASCII file is a text file, EBCDIC file can be transferred if both ends use EBCDIC encoding.
- Image file is the default format for transferring the binary files.
- With ASCII or EBCDIC files one more attribute must be added for defining the printability of the file.
- This attribute is nonprint or TELNET.

8.15.3 Data Structure :

- FTP can use one of the following data structures :
  1. File structure (default)
  2. Record structure and
  3. Page structure.
- File has no structure. It is simply a continuous stream of bytes.
- In the record structure the file is divided into records.
- This data structure is suitable only for the text files.
- In page structure, a file is divided into pages which can be stored or accessed randomly or sequentially.

8.15.4 Transmission Mode :

- FTP uses one of the following modes to transfer a file :
  1. Stream mode
  2. Block mode and
  3. Compressed mode.

1. Stream mode :

- In this mode the data is delivered from FTP to TCP in the form of continuous stream of bytes.
- TCP chops this data into segments of appropriate size.
- Stream mode is the default mode of transmission.

2. Block mode :

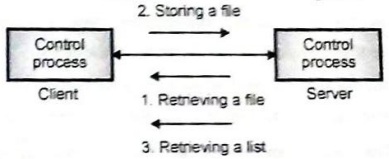
- In this mode, data delivery from FTP to TCP takes place in the form of data blocks.
- Each such block is preceded by a 3 byte header.

3. Compressed mode :

- For big files the data can be compressed. Generally a run length encoding is used for compression.

8.15.5 File Transfer:

- File transfer takes place over the data connection and the commands are sent over the control connection.
- The commands supervise the data transfer.
- But file transfer in FTP means one of the following :
  1. **Retrieving a file** : Server copies a file onto a client.
  2. **Storing of a file** : A file can be copied from client to the server.
  3. A server sends a list of directory or file names to the client. FTP treats such a list of directory also as a file.
- The file transfer has been illustrated in Fig. 8.15.4.



(G-651)**Fig. 8.15.4 : File transfer**

8.16 Remote Login : TELNET :

- The Internet and TCP/IP suite have been designed primarily to provide service to its users.
- The requirements of different users will be of different types and with increase in the number of users, the number of diversified demands will also be very large.
- It is practically impossible to write a specific client - server program for each demand.
- Therefore a general purpose client - server program should be developed which will help a user to access any application on a remote computer.
- That means a user will be allowed to log into a remote computer.
- Two of such general purpose client - server programs which allow remote login are : TELNET and SSH.

8.16.1 TELNET :

- The long form of TELNET is TErminal NETwork. It was proposed by ISO as a standard TCP/IP protocol for a virtual terminal service.
- TELNET enables a user to establish a connection to a remote system.

Concepts related to TELNET :

- Some of the important concepts related to TELNET are as follows :
  1. Time sharing environment.
  2. Login : Local or Remote.
  3. Network Virtual Terminal.

Time Sharing Environment :

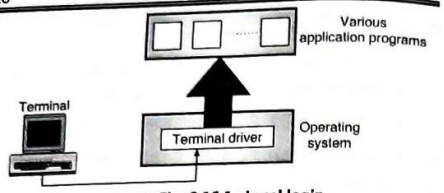
- TELNET was designed during those days when almost all the operating systems were operating on the time - sharing principle.
- In the time sharing environment there is a large central computer which supports all the users.
- All the processing is done by the central computer, and each user feels that it is a dedicated computer.
- The users can access all the common system resources, use all the programs or switch from one program to the other.

Login :

- In a system based on time sharing, every user must have an identification and a password for his authentication.
- Whenever a user wants to access the system he will log into the system with his user id and password.
- The system will check the password to allow only the authorised users to access the resources.
- The logic can be one of the following two types :
  1. Local login.
  2. Remote login.

1. Local login :

- The user login into a local time sharing system is called as local login.
- Fig. 8.16.1 illustrates the principle of local login.

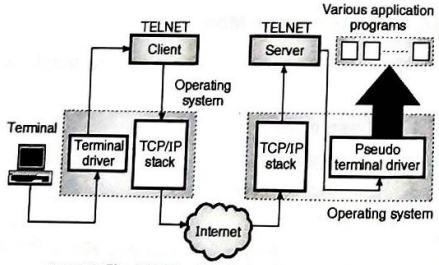


(G-1793) **Fig. 8.16.1 : Local login**

- The local login takes place in a step - by - step manner as follows :
  1. The user types at the keyboard of a terminal.
  2. The terminal driver accepts these keystrokes.
  3. It converts the keystrokes to characters.
  4. It passes the characters to operating system.
  5. The O.S. understands the combination of characters.
  6. It allows access of intended application to the user.

2. Remote Login :

- The user will have to go for the remote login process when he wants to access an application program residing on a remote computer.
- He can do it using the TELNET client and server programs. Fig. 8.16.2 illustrates the principle of remote login.



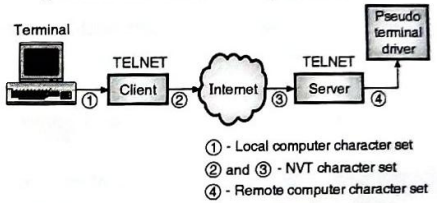
(G-1794) **Fig. 8.16.2 : Principle of remote login**

- Remote login takes place in a step-by-step manner as follows :
  1. The user types at the keyboard of a terminal.
  2. The terminal driver at local O.S. accepts the characters but sends them to TELNET client without interpreting them.
  3. TELNET client converts them into NVT characters. NVT is Network Virtual Terminal. This is a universal character set.

4. NVT characters are delivered to TCP/IP stack (local).
5. The NVT characters travel on the Internet and reach the TCP/IP stack of the remote machine.
6. The NVT characters are applied to the TELNET server which converts them appropriately so that the remote computer can understand them.
7. These characters are applied to a software called pseudo terminal driver.
8. The O.S. at the remote machine then passes the character to the intended application.

8.16.2 Network Virtual Terminal (NVT) :

- NVT character set is a universal interface defined by TELNET in order to ensure that a user can access any remote computer in this world.
- Fig. 8.16.3 illustrates the concept of NVT.



(G-1795) **Fig. 8.16.3 : Concept of NVT**

- The local computer character set is used for the communication between the user terminal and TELNET client.
- Then between the TELNET client and TELNET server the communication takes place using the NVT character set.
- And finally the remote computer character set is used for the communication between the TELNET server and the pseudo terminal driver as shown in Fig. 8.16.3.
- NVT has two sets of characters. One set is for the data and the other set is for control. Both have 8 bit characters.

8.16.3 Security Problems of TELNET :

- TELNET is not a very secured system. It needs username and password for logging in. But it is not enough.
- A snooper software would be enough to capture the login name and password even if they are encrypted.

8.17 Host Configuration : DHCP :

- DHCP (Dynamic host configuration protocol) is the first client server application program that is used after a host is booted.
- Thus it works as a bootstrap when the host is booted and is to be connected to the Internet, but does not know its IP address.
- A computer that makes use of the TCP/IP suite must know its IP address.
- Alongwith its IP address it must also know the following information :
  1. Subnet mask of the computer
  2. IP address of the router, so that it can communicate with other networks.
  3. IP address of the name server so that it can use the names instead of addresses.
- All this information can be saved in a configuration file and accessed by computer when booting takes place.
- This is known as host configuration process.
- But what will happen if the workstation is diskless or the computer is with a disc but it is being booted for the first time.
- If a computer is diskless, then it is possible to store the operating system and networking software in the ROM.
- But this information is not known to the manufacture and therefore cannot be stored in ROM.
- This information is dependent on the configuration of individual machine and it defines which network the machine is connected to.

8.17.1 Previously used Protocols :

- Now a days DHCP has become the formal protocol for host configuration.
- But the two protocols which were used earlier for the same purpose were RARP and BOOTP.
- RARP is Reverse Address Resolution Protocol and BOOTP stands for Bootstrap protocol.

8.17.2 DHCP :

- The Dynamic Host Configuration Protocol (DHCP) was devices by IETF in order to make the configuration automatic.
- Thus DHCP does not require an administrator to add an entry for each computer, to the database that a server uses.
- Instead, in DHCP a mechanism is provided for any computer to join a new network and obtain an IP address automatically with no manual intervention. This is known a plug and play networking.
- Thus DHCP allows the use of computers that, run server software as well as computers that run client software.
- When a computer that runs client software is shifted to a new network, it can use DHCP to obtain configuration information automatically.
- DHCP assigns a permanent address to a nonmobile computer that run server software.
- This address will not change when the computer **reboots.**
- To accommodate both type of computers, DHCP makes use of a client server approach.
- When a computer boots, it will broadcasts a DHCP Request. In response a server sends a DHCP Replay.
- An administrator can configure a DHCP server to have two types of addresses.
- First is the permanent address that are assigned to server computers, and second type is a pool of addresses which can be assigned on the basis of demand, when a computer boots and sends a request to DHCP.
- The DHCP find the configuration information by accessing its database If the database contains a specific entry for the computer then the server returns the information from the entry.
- However if there is no such entry exists for the computer, then the server chooses the next IP address from the pool and assigns it to the computer.

What is DHCP :

- DHCP, as the name suggests, is a protocol used for dynamically configuring the hosts on a network, such as workstations, personal computers and printers.
- DHCP can help in assigning various types of information such as routing information, directory-services information and default web server and mail servers.
- However, the most important and commonly used information for which DHCP is used is the IP address and subnet mask information.
- DHCP was primarily designed for managing the network and the clients automatically.
- With DHCP, it is not necessary to configure the network and client information manually for individual hosts.
- In addition, DHCP can coexist with statically configured hosts with fixed IP addresses.
- DHCP can also carry out the allocation of certain configuration information to a host on a permanent basis.
- This protocol provides a four point information (IP address, subnet mask, IP address of router, IP address of name server) to a diskless computer or to a computer which is booted for the first time.
- It is a client / server protocol which is backward compatible to the BOOTP.

8.17.3 Advantages of DHCP :

- The use of DHCP on a network offers the following advantages :

1. It sets free the network administrator from the duties of setting up the configuration information, such as the IP address, the subnet mask, and the routing tables, manually. The DHCP simplifies network administration by doing these tasks automatically.
2. Avoids this and the sometimes the same IP address is assigned to two different hosts. The DHCP avoids this and the consequent malfunctioning of both the hosts from happening.

3. If the DHCP was not used, then the movement of computers from one network to another requires must be reconfigured. With DHCP, you can move the computers to different subnets or networks without the need to reconfigure them. In such situations, DHCP takes care of IP address assignment and other configuration details.
4. Mobile computers, such as laptops and palmtops, can easily get connected to different networks. They don't require reconfiguration any more as they get their configuration information from the DHCP server.
5. DHCP allocates IP addresses from a pool of IP addresses. In addition, when a computer gets disconnected, its released IP address is returned to the resource pool. Therefore, the possibility of having unused IP addresses are minimized.

8.17.4 Components of DHCP :

- The use of DHCP on a network requires the following three components :

1. DHCP server :

- It assigns the IP address and other information to the clients when they request for the information.

2. DHCP client :

- It communicates with the DHCP server to get the desired information regarding its configuration.
- This communication can take place when the computer starts.
- The user of the DHCP client can also initiate a DHCP client request to the DHCP server to renew its information.

3. DHCP relay agent :

- It is used to relay (forward) client requests to the DHCP server.
- This is required when the DHCP server is yet to assign the client an IP address.
- Without an IP address, a client cannot use IP routing on its own.
- A DHCP relay agent helps the client to communicate with the DHCP server when the client does not have an IP address.

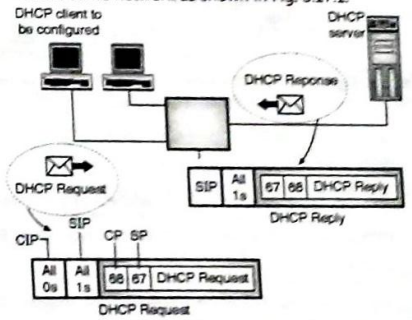
- When a client starts, it has an IP address of 0.0.0.0. It sends a broadcast message containing its MAC address and the computer name.
- In response the DHCP server sends an offer message that contains the MAC address of the client, the IP address offered to that client, the lease period for which the IP address will remain valid and its own IP address.
- The lease period is the time duration for which a client can use the IP address that has been assigned to it by the DHCP server.
- You can configure a DHCP server to set the lease time.
- When the client receives the IP address, it accepts the offer and then broadcasts the message that it has accepted the offer.

8.17.5 DHCP Operation :

- We will discuss the DHCP operation under two different operating conditions :
  1. DHCP client and server on the same network.
  2. DHCP client and server on different networks.

Operation on the same network :

- This situation is not a very common one. But sometimes the DHCP client and server happen to be on the same network as shown in Fig. 8.17.1.



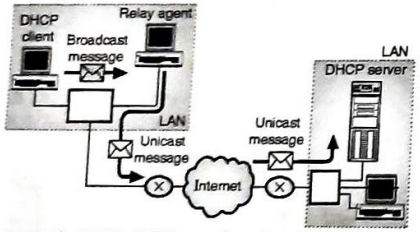
(G-1789) Fig. 8.17.1 : Operation of DHCP when client and server are on the same network

- The operation takes place as follows :
  1. The DHCP server sends a passive open command on port 67 of UDP and waits for clients response.

2. The DHCP client sends an active open command on port 68 of UDP. This message is encapsulated in the UDP datagram with port 67 as destination port and port 68 as the source port. The UDP datagram is then encapsulated in an IP datagram. Note that the client at this time does not know its own IP address (i.e. the source address) and the server's IP address (destination address). Therefore the client uses an all zero address as source address and an all one address as destination address.
3. The server responds to this message by sending either a broadcast or a unicast message using port 67. It uses port 68 as the destination port. Broadcast address is used only for those system which do not allow the bypassing of ARP.

8.17.6 DHCP Operation on Different Networks :

- In this situation the DHCP client and server are on two entirely different networks, as shown in Fig. 8.17.2.



(G-1790) Fig. 8.17.2 : DHCP operation when client and server are on different networks

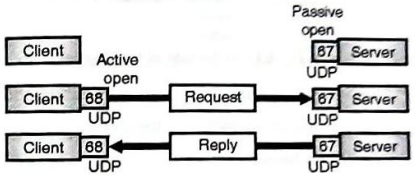
- In this situation a problem arises due to the broadcast nature of DHCP request.
- The client does not know the IP address of the server.
- Hence the DHCP request is a broadcast type (all 1s IP address).
- Any server does not allow the broadcast request to pass through it. So this request cannot reach the DHCP server.
- In order to solve this problem we can configure one of the hosts or router to operate as a relay agent as shown in Fig. 8.17.2.
- The relay agent knows the unicast address of the DHCP server.
- The relay will look for the broadcast request on port 67.
- As soon as it receives the broadcast request message, it encapsulates this message in a unicast datagram and sends it to the DHCP server.
- Such a unicast message is allowed to pass through by any router. Thus the request message reaches the DHCP server.
- The DHCP server sends its reply to the relay agent which in turn sends it to the DHCP client.

Note : In Fig. 8.17.2 only the message between the relay agent and client is it broadcast type. All the other messages are unicast types.

8.17.7 UDP Ports :

- The interaction between a client and DHCP server has been shown in Fig. 8.17.3.
- The well known port 67 is used by the server, which is normal.
- But the client uses the well known port 68, which is not normal. It is unusual.
- Why does a client choose the well known port 68 rather than an ephemeral port ?
- The answer is for prevention of a problem when the reply from the server to client is of **broadcast** type.
- In order to understand the exact nature of the problem, let us assume that an **ephemeral port** is used instead of the well known port 68 and study its effect.
- Suppose host A on a network is using a DHCP client.
- It is using the ephemeral port say 2017 which we have chosen randomly.
- On the same network, there is another host B, which is using a DAYTIME client on ephemeral port 2017 which is accidentally the same.
- In this situation, the DHCP server sends a broadcast reply message with the destination port number 2017 and broadcast IP address $FFFFFFFF_{16}$.

- Every host has to open a packet which carries this destination IP address.
- Host A would find a message from an application program on ephemeral port 2017.
- Thus the DHCP client receives a **correct message** but the DAYTIME client receives an **incorrect message.**
- This confusion takes place due to the process of demultiplexing which is based on the **socket address.**
- Remember that a socket address is the combination of IP address and port number and both are same in this case.



(G-1994) Fig. 8.17.3 : Use of UDP ports

- If a well known port (less than 1024) is used then the use of same two destination port numbers would be prevented.
- It would not be possible for host B to select port 68 as the ephemeral port due to the fact that ephemeral port numbers are greater than 1023.
- The final question is what happens if host B is also running the DHCP client ?
- The answer is that because of the same socket address, both the clients will receive the message.
- In order to handle such a situation, **the third identification number** is used to differentiate the clients.
- In DHCP this another number is called as the **transactional ID** and for each DHCP connection, it is chosen randomly.
- It is almost impossible that both the host choose the same transactional ID.

8.17.8 Using TFTP :

- Note that all the information needed by a client for booting purpose is not sent by the server.
- The server, in its reply message will define the **pathname** of a file in which all the booting information needed for the client is sent.
- The client can then use a TFTP message that is encapsulated in a UDP user datagram, to obtain the remaining necessary information.

8.17.9 Error Control :

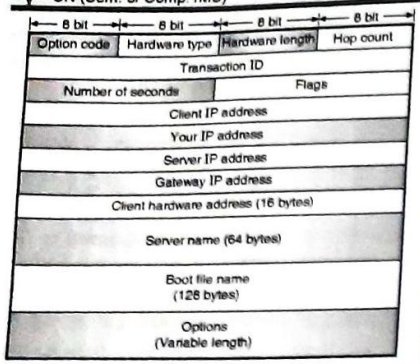
- DHCP can use either UDP (as discussed) or TFTP. Note that UDP does not provide any error control.
- Then what should be done if a request is lost or damaged ? OR if the reply is damaged ?
- As UDP does not provide any error control, the DHCP should provide it.
- Two strategies could be used to achieve the goal of error control :
  1. Ask UDP to use checksum. The UDP has an option of using the checksum.
  2. Ask DHCP client to use timers alongwith the retransmission policy if DHCP request or reply gets damaged or lost.

8.17.10 Optimizations in DHCP :

- The DHCP protocol has following steps :
- The first step is that a computer **broadcasts** a DHCP discover message in order to find DHCP server, and the other step is that the computer selects one of the available DHCP servers that responds to its message and sends a request to that server.
- To avoid a situation in which a computer follows both steps each time its boots or each time it needs to extend the lease, DHCP uses caching.
- When a computer discovers a DHCP server, the computer saves the address of that server in a cache on **permanent** storage (e.g. a disk file).
- Similarly, once an IP address has been allotted to it the computer saves the IP address in a cache. When a computer reboots, it uses the cached information to revalidate its former address.
- Doing so saves time and reduce network traffic.

8.17.11 Packet Format :

- The format of a DHCP packet has been shown in Fig. 8.17.4.



(G-1995) **Fig. 8.17.4 : DHCP packet format**

- Let us describe each field in the DHCP packet.

1. Operation code :

- This is an 8 bit field which is used to define the type of DHCP packet.
- If this field contains (1) then the packet is **request type** and if this field contains (2) then the packet is **reply type.**

2. Hardware type :

- This 8-bit field is used to define the type of physical network.
- An integer has been assigned to each type of network e.g. the value of this field is 1 for Ethernet.

3. Hardware length :

- This is an 8-bit field which is used for defining the length of the physical address in bytes.
- The value of this field is 6 for Ethernet because the physical address of Ethernet is 6 byte long.

4. Hop count :

- This is an 8-bit field which is used for define the maximum number of hops a packet can travel.

5. Transaction ID :

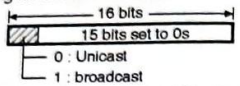
- This is a 32-bit or 4-byte long field which carries an integer in it.
- The contents of this field are known as transaction **identification** and it is set by the client.
- This field is used for matching a reply with the request.
- The same value is returned by the server in its reply packet.

6. Number of seconds :

- This is a 16-bit field which is used to indicate the amount of time (in seconds) elapsed from the instant at which the client started to boot.

7. Flag :

- This is a 16-bit long field, as shown in Fig. 8.17.5. Out of these 16 bits, only the leftmost bit is used and the remaining 15 bits are set to 0s.



(G-1996) **Fig. 8.17.5 : Format of the flagfield**

- The leftmost bit is used to specify a forced broadcast reply (instead of unicast) from the server.

8. Client IP address :

- This 4-byte long field is used to carry the client IP address.
- A "0" in this field indicates that the client does not have this information.

8. Your IP address :

- This is also a 4-byte long field which is used to carry the clients IP address.
- This address is requested by the client and filled by the server in the reply message.

10. Server IP address :

- This is also a 4-byte long field which contains the IP address of the server.
- This address is sent by the server in the reply message.

11. Gateway IP address :

- This is a 4-byte or 32 bit long field that contains the IP address of a router which is filled in the reply message by the server.

12. Client hardware address :

- This is a 16-byte field which contains the physical address of the client.

13. Server name :

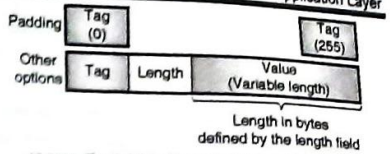
- This is a 64 byte long field which is filled on the optional basis by the server in a reply packet.
- This field consists of a null terminated string containing the domain name of the server.
- If no information about the server name is to be given, then the server should fill up this field with all zeros.

14. Boot filename :

- This is a 128-byte field which contains a null terminated string consisting of full pathname of the boot file.
- This path can be used by the client in order to obtain additional information about booting.
- This field is filled by the server in the reply message on the optional basis.
- If the server does not want to fill data in this field, then the entire field should be filled up with 0s.

15. Options :

- This is a 64-byte field which can be used for a dual purpose as follows :
  1. It is used to carry some additional information such as default router address or network mask.
  2. Or it is used to carry some specific information about the vendor.
- It is important to note that, the **options** field is used only in the **reply message.**
- The server makes use of a number called **magic cookie.**
- After finishing reading of the message the client searches for the magic cookie.
- If it is present, then the next 60 bytes data will correspond to **options.**
- Fig. 8.17.6 shows the format of the option. It consists of three fields as follows : a 1-byte **tag** field, a 1-byte **length** field and a variable length **value** field.



(G-1997) **Fig. 8.17.6 : Format of the options field**

- The function of the length field is to specify the length of the variable length **value** field and not of the whole option.

Review Questions

Q. 1 Explain in brief about the application layer.

Q. 2 Write a short note on providing services.

Q. 3 Explain about the standard and nonstandard protocols at the application layer.

Q. 4 Explain in brief client-server paradigm.

Q. 5 State the problems and applications of client-server paradigm.

Q. 6 Explain the P2P paradigm.

Q. 7 State the merits, demerits and applications of P2P paradigm.

Q. 8 Explain the term API and state its types.

Q. 9 Define a socket and state its role.

Q. 10 Draw and explain the structure of www.

Q. 11 Explain the non-persistent and persistent connections in HTTP.

Q. 12 Write a note on : HTTP messages.

Q. 13 What is FTP ? Explain the communication in FTP.

Q. 14 Write a note on E-mail.

Q. 15 Compare SMTP and HTTP.

Q. 16 Write a note on message access agents.

Q. 17 Briefly discuss the following terms, emphasis more on implementation details :

(a) DNS (b) Mail server