UNIT 1 INTRODUCTION TO COMPUTER NETWORKS

Structure **Page Nos.**

1.0 INTRODUCTION

These days, practically every business, no matter how small uses computers to handle various transactions and as business grows, they often need several people to input and process data simultaneously and in order to achieve this, the earlier model of a single computer serving all the organisations computational needs has been replaced by a model in which a number of separate but interconnected computers do the job and this model is known as a Computer Network. By linking individual computers over a network their productivity has been increased enormously.

A most distinguishing characteristic of a general computer network is that data can enter or leave at any point and can be processed at any workstation. For example: A printer can be controlled from any word processor at any computer on the network. This is an introductory unit where, you will be learning about the basic concepts regarding Computer Networks. Here, you will learn about Networks, different types of Networks, their applications, Network topology, Network protocols, OSI Reference Model, TCP/IP Reference Model. We shall also examine some of the popular computer networks like Novell network, ARPANET, Internet, and ATM networks.

Towards the end of this unit the concept of Delays in computer networks is also discussed.

1.1 OBLECTIVES

After going through this unit, you should be able to:

- understand the concept of computer networks;
- differentiate between different types of computer networks;
- understand the different application of networks;
- compare the different network topologies;
- signify the importance of network protocols;
- know the importance of using networked system;
- understand the layered organisation and structuring of computer networks using OSI and TCP/IP reference model;
- have a broad idea about some of the popular networks like Novell network, ARPANET, INTERNET, ATM etc., and
- understand the concept of delays.

1.2 WHAT IS A COMPUTER NETWORK?

Figure 1: A computer-networked environment

A Computer network consists of two or more autonomous computers that are linked (connected) together in order to:

- Share resources (files, printers, modems, fax machines).
- Share Application software like MS Office.
- Allow Electronic communication.
- Increase productivity (makes it easier to share data amongst users).

Figure 1 shows people working in an networked environment. The Computers on a network may be linked through Cables, telephones lines, radio waves, satellites etc.

A Computer network includes, the network operating system in the client and server machines, the cables, which connect different computers and all supporting hardware

in between such as bridges, routers and switches. In wireless systems, antennas and towers are also part of the network.

Computer networks are generally classified according to their structure and the area they are localised in as:

- **Local Area Network (LAN):** The network that spans a relatively small area that is, in the single building or campus is known as LAN.
- **Metropolitan Area Network (MAN):** The type of computer network that is, designed for a city or town is known as MAN.
- **Wide Area Network (WAN):** A network that covers a large geographical area and covers different cities, states and sometimes even countries, is known as WAN.

The additional characteristics that are also used to categorise different types of networks are:

- **Topology:** Topology is the graphical arrangement of computer systems in a network. Common topologies include a bus, star, ring, and mesh.
- **Protocol:** The protocol defines a common set of rules which are used by computers on the network that communicate between hardware and software entities. One of the most popular protocols for LANs is the Ethernet. Another popular LAN protocol for PCs is the token-ring network.
- **Architecture:** Networks can be broadly classified as using either a peer-to-peer or client/server architecture.

1.3 NETWORK GOALS AND MOTIVATIONS

Before designing a computer network we should see that the designed network fulfils the basic goals. We have seen that a computer network should satisfy a broad range of purposes and should meet various requirements. One of the main goals of a computer network is to enable its users to share resources, to provide low cost facilities and easy addition of new processing services. The computer network thus, creates a global environment for its users and computers.

Some of the basic goals that a Computer network should satisfy are:

- Cost reduction by sharing hardware and software resources.
- Provide high reliability by having multiple sources of supply.
- Provide an efficient means of transport for large volumes of data among various locations (High throughput).
- Provide inter-process communication among users and processors.
- Reduction in delay driving data transport.
- Increase productivity by making it easier to share data amongst users.
- Repairs, upgrades, expansions, and changes to the network should be performed with minimal impact on the majority of network users.
- Standards and protocols should be supported to allow many types of equipment from different vendors to share the network (Interoperatability).
- Provide centralised/distributed management and allocation of network resources like host processors, transmission facilities etc.

1.4 CLASSIFICATION OF NETWORKS

Depending on the transmission technology i.e., whether the network contains switching elements or not, we have two types of networks:

- Broadcast networks.
- Point-to-point or Switched networks.

1.4.1 Broadcast Networks

Broadcast networks have a single communication channel that is shared by all the machines on the network. In this type of network, short messages sent by any machine are received by all the machines on the network. The packet contains an address field, which specifies for whom the packet is intended. All the machines, upon receiving a packet check for the address field, if the packet is intended for itself, it processes it and if not the packet is just ignored.

Using Broadcast networks, we can generally address a packet to all destinations (machines) by using a special code in the address field. Such packets are received and processed by all machines on the network. This mode of operation is known as "Broadcasting". Some Broadcast networks also support transmission to a subset of machines and this is known as "Multicasting". One possible way to achieve Multicasting is to reserve one bit to indicate multicasting and the remaining (n-1) address bits contain group number. Each machine can subscribe to any or all of the groups.

Broadcast networks are easily configured for geographically localised networks. Broadcast networks may be Static or dynamic, depending on how the channel is allocated.

In Static allocation, time is divided into discrete intervals and using round robin method, each machine is allowed to broadcast only when its time slot comes up. This method is inefficient because the channel capacity is wasted when a machine has nothing to broadcast during its allocated slot.

Dynamic allocation may be centralised or decentralised. In centralised allocation method, there is a single entity, for example, a bus arbitration unit which determine who goes next and this is achieved by using some internal algorithm. In Decentralised channel allocation method, there is no central entity, here, each machine decides for itself whether or not to transmit.

The different types of Broadcast networks are:

1) Packet Radio Networks.
2) Satellite Networks.
3) Local Area Networks.

Packet Radio broadcasting differs from **satellite network** broadcasting in several ways. In particular stations have limited range introducing the need for radio repeaters, which in turn affects the routing, and acknowledges schemes. Also the propagation delay is much less than for satellite broadcasting.

LAN (Local Area Network)

Local Area Network is a computer network that spans over a relatively small area. Most LANs are confined to a single building or group of buildings within a campus. However, one LAN can be connected to other LANs over any distance via telephone

lines and radio waves. A system of LANs connected in this way is called a wide-area network (WAN).

Most LANs connect workstations and personal computers. Each node (individual computer) in a LAN has its own CPU with which it executes programs, but it is also able to access data and devices anywhere on the LAN. This means that many users can share data as well as expensive devices, such as laser printers, fax machines etc. Users can also use the LAN to communicate with each other, by sending e-mail or engaging in chat sessions. There are many different types of LANs, Ethernets being the most common for PCs.

The following characteristics differentiate one LAN from another:

- **Topology:** The geometric arrangement of devices on the network. For example, devices can be arranged in a ring or in a straight line.
- **Protocols:** The rules and encoding specifications for sending data. The protocols also determine whether the network uses peer-to-peer or client/server architecture.
- **Media:** Devices can be connected by twisted-pair wire, coaxial cables, or fiber optic cables. Some networks communicate via radio waves hence, do not use any connecting media.

LANs are capable of transmitting data at very fast rates, much faster than data can be transmitted over a telephone line; but the distances are limited, and there is also a limit on the number of computers that can be attached to a single LAN.

The typical characteristics of a LAN are:

- Confined to small areas i.e., it connects several devices over a distance of 5 to 10 km.
- High speed.
- Most inexpensive equipment.
- Low error rates.
- Data and hardware sharing between users owned by the user.
- Operates at speeds ranging from 10Mbps to 100Mbps. Now a days 1000 Mbps are available.

1.4.2 Point-to-Point or Switched Networks

Point–to-point or switched, networks are those in which there are many connections between individual pairs of machines. In these networks, when a packet travels from source to destination it may have to first visit one or more intermediate machines. Routing algorithms play an important role in Point-to-point or Switched networks because often multiple routes of different lengths are available.

An example of switched network is the international dial-up telephone system.

The different types of Point- to-point or Switched networks are:

- Circuit Switched Networks.
- Packet Switched Networks.

In Switched network, the temporary connection is established from one point to another for either the duration of the session (circuit switching) or for the transmission of one or more packets of data (packet switching).

Circuit Switched Networks

Circuit Switched networks use a networking technology that provides a temporary, but dedicated connection between two stations no matter how many switching devices are used in the data transfer route. Circuit switching was originally developed for the analog based telephone system in order to guarantee steady and consistent service for two people engaged in a phone conversation. Analog circuit switching has given way to digital circuit switching, and the digital counterpart still maintains the connection until broken (one side hangs up). This means bandwidth is continuously reserved and "silence is transmitted" just the same as digital audio in voice conversation.

Packet Switched Networks

Packet switched Networks use a networking technology that breaks up a message into smaller packets for transmission and switches them to their required destination. Unlike circuit switching, which requires a constant point-to-point circuit to be established, each packet in a packet- switched network contains a destination address. Thus, all packets in a single message do not have to travel the same path. They can be dynamically routed over the network as lines become available or unavailable. The destination computer reassembles the packets back into their proper sequence.

Packet switching efficiently handles messages of different lengths and priorities. By accounting for packets sent, a public network can charge customers for only the data they transmit. Packet switching has been widely used for data, but not for real-time voice and video. However, this is beginning to change. IP and ATM technologies are expected to enable packet switching to be used for everything.

The first international standard for wide area packet switching networks was X.25, which was defined when all circuits were digited and susceptible to noise. Subsequent technologies, such as frame relay and SMDS were designed for today's almost-error-free digital lines.

ATM uses a cell-switching technology that provides the bandwidth sharing efficiency of packet switching with the guaranteed bandwidth of circuit switching.

Higher-level protocols, such as TCP/IP, IPX/SPX and NetBIOS, are also packet based and are designed to ride over packet-switched topologies.

Public packet switching networks may provide value added services, such as protocol conversion and electronic mail.

☞ Check Your Progress 1

1) Explain the difference between Client/Server and Peer-to-peer architecture.

……

……

……

2) List the important aspects that should be kept in mind while designing a network?

……

……

……

3) Write briefly about the areas where networks are used?

……

……

……

4) Differentiate between Broadcast and point-to-point networks.

……

……

……

1.5 NETWORK TOPOLOGY

Topology refers to the shape of a network, or the network's layout. How different nodes in a network are connected to each other and how they communicate with each other is determined by the network's topology. Topologies are either *physical* or *logical*.

Some of the most common network topologies are:

- Bus topology
- Star topology
- Ring topology
- Tree topology
- Mesh topology
- Cellular topology.

The parameters that are to be considered while selecting a physical topology are:

- Ease of installation.
- Ease of reconfiguration.
- Ease of troubleshooting.

1.5.1 Bus Topology

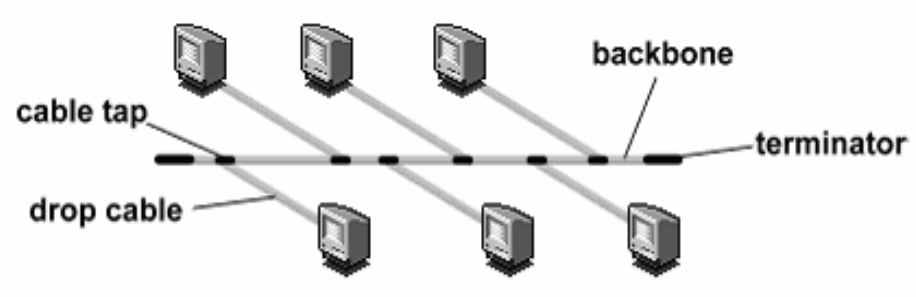

Figure 2: Bus topology

In Bus topology, all devices are connected to a central cable, called the bus or backbone. The bus topology connects workstations using a single cable. Each workstation is connected to the next workstation in a point-to-point fashion. All workstations connect to the same cable. *Figure 2* shows computers connected using Bus Topology.

In this type of topology, if one workstation goes faulty all workstations may be affected as all workstations share the same cable for the sending and receiving of information. The cabling cost of bus systems is the least of all the different topologies. Each end of the cable is terminated using a special terminator.

The common implementation of this topology is Ethernet. Here, message transmitted by one workstation is heard by all the other workstations.

Advantages of Bus Topology

- Installation is easy and cheap when compared to other topologies.
- Connections are simple and this topology is easy to use.
- Less cabling is required.

Disadvantages of Bus Topology

- Used only in comparatively small networks.
- As all computers share the same bus, the performance of the network deteriorates when we increase the number of computers beyond a certain limit.
- Fault identification is difficult.
- A single fault in the cable stops all transmission.

1.5.2 Star Topology

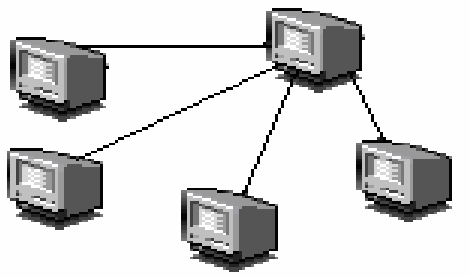

Figure 3: Star topology

Star topology uses a central hub through which, all components are connected. In a Star topology, the central hub is the host computer, and at the end of each connection is a terminal as shown in *Figure 3.*

Nodes communicate across the network by passing data through the hub. A star network uses a significant amount of cable as each terminal is wired back to the central hub, even if two terminals are side by side but several hundred meters away from the host. The central hub makes all routing decisions, and all other workstations can be simple.

An advantage of the star topology is, that failure, in one of the terminals does not affect any other terminal; however, failure of the central hub affects all terminals. This type of topology is frequently used to connect terminals to a large time-sharing host computer.

Advantages of Star Topology

- Installation and configuration of network is easy.
- Less expensive when compared to mesh topology.
- Faults in the network can be easily traced.

- Expansion and modification of star network is easy.
- Single computer failure does not affect the network.
- Supports multiple cable types like shielded twisted pair cable, unshielded twisted pair cable, ordinary telephone cable etc.

Disadvantages of Star Topology

- Failure in the central hub brings the entire network to a halt.
- More cabling is required in comparison to tree or bus topology because each node is connected to the central hub.

1.5.3 Ring Topology

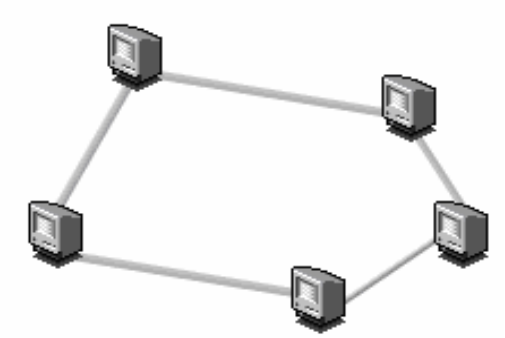

Figure 4: Ring topology

In Ring Topology all devices are connected to one another in the shape of a closed loop, so that each device is connected directly to two other devices, one on either side of it, i.e., the ring topology connects workstations in a closed loop, which is depicted in *Figure 4* Each terminal is connected to two other terminals (the next and the previous), with the last terminal being connected to the first. Data is transmitted around the ring in one direction only; each station passing on the data to the next station till it reaches its destination.

Information travels around the ring from one workstation to the next. Each packet of data sent on the ring is prefixed by the address of the station to which it is being sent. When a packet of data arrives, the workstation checks to see if the packet address is the same as its own, if it is, it grabs the data in the packet. If the packet does not belong to it, it sends the packet to the next workstation in the ring.

Faulty workstations can be isolated from the ring. When the workstation is powered on, it connects itself to the ring. When power is off, it disconnects itself from the ring and allows the information to bypass the workstation.

The common implementation of this topology is token ring. A break in the ring causes the entire network to fail. Individual workstations can be isolated from the ring.

Advantages of Ring Topology

- Easy to install and modify the network.
- Fault isolation is simplified.
- Unlike Bus topology, there is no signal loss in Ring topology because the tokens are data packets that are re-generated at each node.

Disadvantages of Ring Topology

- Adding or removing computers disrupts the entire network.

- A break in the ring can stop the transmission in the entire network.
- Finding fault is difficult.
- Expensive when compared to other topologies.

1.5.4 Tree Topology

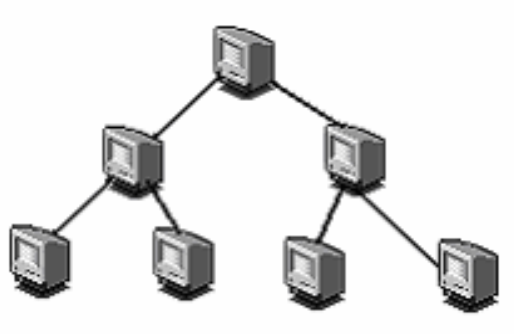

Figure 5: Tree topology

Tree topology is a LAN topology in which only one route exists between any two nodes on the network. The pattern of connection resembles a tree in which all branches spring from one root. *Figure 5* shows computers connected using Tree Topology.

Tree topology is a hybrid topology, it is similar to the star topology but the nodes are connected to the secondary hub, which in turn is connected to the central hub. In this topology groups of star-configured networks are connected to a linear bus backbone.

Advantages of Tree Topology

- Installation and configuration of network is easy.
- Less expensive when compared to mesh topology.
- Faults in the network can be detected traced.
- The addition of the secondary hub allows more devices to be attached to the central hub.
- Supports multiple cable types like shielded twisted pair cable, unshielded twisted pair cable, ordinary telephone cable etc.

Disadvantages of Tree Topology

- Failure in the central hub brings the entire network to a halt.
- More cabling is required when compared to bus topology because each node is connected to the central hub.

1.5.5 Mesh Topology

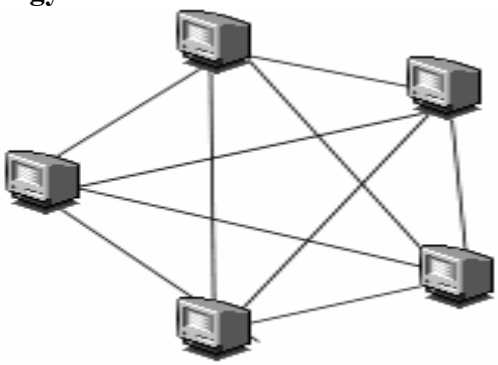

Figure 6: Mesh Topology

Devices are connected with many redundant interconnections between network nodes. In a well-connected topology, every node has a connection to every other node in the

network. The cable requirements are high, but there are redundant paths built in. Failure in one of the computers does not cause the network to break down, as they have alternative paths to other computers.

Mesh topologies are used in critical connection of host computers (typically telephone exchanges). Alternate paths allow each computer to balance the load to other computer systems in the network by using more than one of the connection paths available. A fully connected mesh network therefore has n (n-1)/2 physical channels to link n devices. To accommodate these, every device on the network must have (n-1) input/output ports.

Advantages of Mesh Topology

- Use of dedicated links eliminates traffic problems.
- Failure in one of the computers does not affect the entire network.
- Point-to-point link makes fault isolation easy.
- It is robust.
- Privacy between computers is maintained as messages travel along dedicated path.

Disadvantages of Mesh Topology

- The amount of cabling required is high.
- A large number of I/O (input/output) ports are required.

1.5.6 Cellular Topology

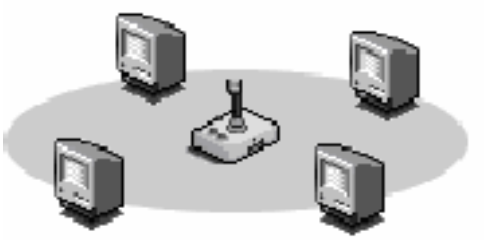

Figure 7: Cellular topology

Cellular topology, divides the area being serviced into cells. In wireless media each point transmits in a certain geographical area called a cell, each cell represents a portion of the total network area. *Figure 7* shows computers using Cellular Topology. Devices that are present within the cell, communicate through a central hub. Hubs in different cells are interconnected and hubs are responsible for routing data across the network. They provide a complete network infrastructure.
Cellular topology is applicable only in case of wireless media that does not require cable connection.

Advantages of Cellular Topology

- If the hubs maintain a point-to-point link with devices, trouble shooting is easy.
- Hub-to-hub fault tracking is more complicated, but allows simple fault isolation.

Disadvantages of Cellular Topology

- When a hub fails, all devices serviced by the hub lose service (are affected).

☞ Check Your Progress 2

1) List the importance of using computer networks in Airline and Railway reservation systems?

……

……

……

2) What are the various types of networks?

……

……

……

3) Compare Tree topology with Star topology.

……

……

……

4) Suppose we have to add new nodes to the network, then which is the best suited topology and why?

……

……

……

1.6 APPLICATIONS OF NETWORK

Computer networks are used as a highly reliable medium for exchange of information. Using a computer network we can do virtually everything that a Mainframe or a Minicomputer can do, but at a much lower cost.

There are numerous applications of computer networks some of them are:

- Share resources and information.
- Access to remote information.
- Person-to-person communication.
- Interactive entertainment.

Share Resources and Information

Using a Computer network we can share expensive resources such as laser printers, CD-ROM Drives, Fax machines etc. We can share information and many persons can work together on projects and tasks that require co-ordination and communication, even though these users may not be physically close.

Access to Remote Information

Access to remote information involves interaction between a person and a remote database. Financial Institutions allow access to their information so that people can pay their bills, handle their investments and manage their bank accounts electronically. Online shopping also allows people, access to product information before purchasing the product.

These days, many newspapers and digital libraries are available online and allow users to access news and information which is of interest to them. Another application is the World Wide Web, which contains information about a wide variety of subjects like health, sports, science, recreation, history, government etc.

Person-to-person Communication

Person-to-person communication is mostly carried out using e-mail (Electronic mail) Electronic mail is a simple, yet potent facility. E-mail is more valuable and useful than the telephone because by using e-mail we can convey information that is difficult or impossible to read over the telephone, like reports, tables, charts, images etc.
Using a computer network, it is also possible to organise virtual meeting among people who are far away from each other and this is called video conferencing. Virtual meetings have other applications such as in Distance education, getting medical opinions from distant specialists (remote diagnosis) etc. Multimedia communication can also be used for tele training.

Interactive Entertainment

Computer Networks such as Internet offer a wide variety of entertainment, there are many companies online, which offer video-on-demand. A large variety of multi-person real-time simulation games, like hide-and seek in a virtual dungeon and flight simulators with the players in one team trying to shoot down the players in the opposite team and many such games are available on-line.

1.7 NETWORKING MODEL

Most of the networks today are organised as a series of stacked layers with each layer stacked over another layer below it. This is done in order to divide the workload and to simplify the systems design. The architecture is considered scalable if it is able to accommodate a number of layers in either large or small scales. For example, a computer that runs an Internet application may require all of the layers that were defined for the architectural model. Similarly, a computer that acts as a router may not need all these layers. Systems design is furthermore simplified because with a layered architecture, the design has to only concern the layer in question and not worry about the architecture in a macro sense.

The depth and functionality of this stack differs from network to network. However, regardless of the differences among all networks, the purpose of each layer is to provide certain services (job responsibilities) to the layer above it, shielding the upper layers from the intricate details of how the services offered are implemented.

Every computer in a network possesses within it a generic stack. A *logical* communication may exist between any two computers through the layers of the same "level". Layer-n on one computer may converse with layer-n on another computer. There are rules and conventions used in the communication at any given layers, which are known collectively as the layer-n *protocol* for the nth layer.

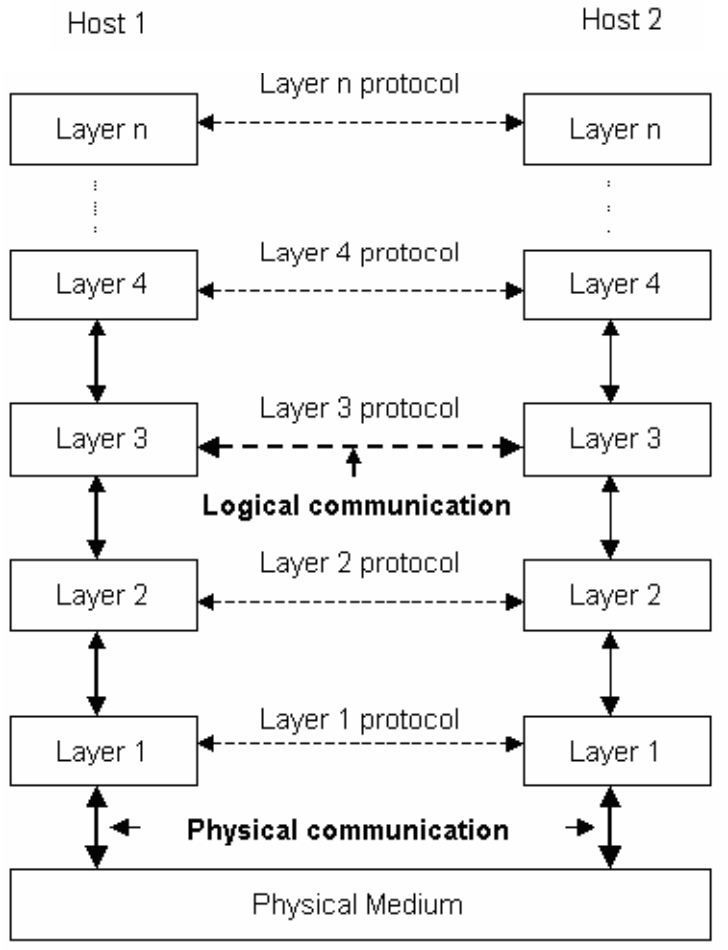

Figure 8: Layered network architecture

Data are not directly transferred from layer-n on one computer to layer-n on another computer. Rather, each layer passes data and control information to the layer directly below until the lowest layer is reached. Below layer-1 (the bottom layer), is the physical medium (the hardware) through which the actual transaction takes place. In *Figure 8* logical communication is shown by a broken-line arrow and physical communication by a solid-line arrow.

Between every pair of adjacent layers is an interface. The interface is a specification that determines how the data should be passed between the layers. It defines what primitive operations and services the lower layer should offer to the upper layer. One of the most important considerations when designing a network is to design clean-cut interfaces between the layers. To create such an interface between the layers would require each layer to perform a specific collection of well-understood functions. A clean-cut interface makes it easier to replace the implementation of one layer with another implementation because all that is required of the new implementation is that, it offers, exactly the same set of services to its neighbouring layer above as the old implementation did.

1.7.1 OSI Reference Model

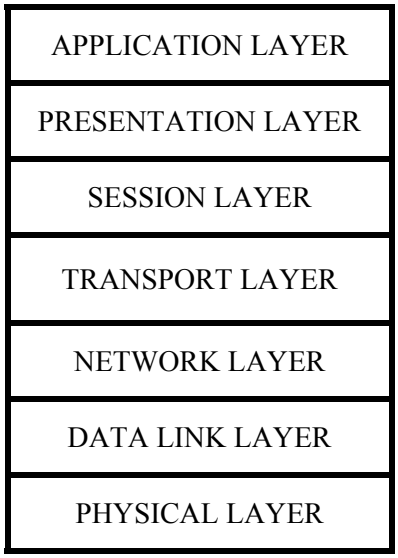

Figure 9: Layers of OSI reference model

The Open System Interconnection (OSI) model is a set of protocols that attempt to define and standardise the data communications process; we can say that it is a concept that describes how data communications should take place.

The OSI model was set by the International Standards Organisation (ISO) in 1984, and it is now considered the primary architectural model for inter-computer communications. The OSI model has the support of most major computer and network vendors, many large customers, and most governments in different countries.

The Open Systems Interconnection (OSI) reference model describes how information from a software application in one computer moves through a network medium to a software application in another computer. The OSI reference model is a conceptual model composed of seven layers as shown in *Figure 9* each specifying particular network functions and into these layers are fitted the protocol standards developed by the ISO and other standards bodies. The OSI model divides the tasks involved with moving information between networked computers into seven smaller, more manageable task groups. A task or group of tasks is then assigned to each of the seven OSI layers. Each layer is reasonably self-contained so that the tasks assigned to each layer can be implemented independently. This enables the solutions offered by one layer to be updated without affecting the other layers.

The OSI model is modular. Each successive layer of the OSI model works with the one above and below it.

Although, each layer of the OSI model provides its own set of functions, it is possible to group the layers into two distinct categories. The first four layers i.e., physical, data link, network, and transport layer provide the end-to-end services necessary for the transfer of data between two systems. These layers provide the protocols associated with the communications network used to link two computers together. Together, these are communication oriented.

The top three layers i.e., the application, presentation, and session layers provide the application services required for the exchange of information. That is, they allow two applications, each running on a different node of the network to interact with each other through the services provided by their respective operating systems. Together, these are data processing oriented.

The following are the seven layers of the Open System Interconnection (OSI) reference model:

- Layer 7 — Application layer
- Layer 6 — Presentation layer
- Layer 5 — Session layer
- Layer 4 — Transport layer
- Layer 3 — Network layer
- Layer 2 — Data Link layer
- Layer 1 — Physical layer

Application layer (Layer 7)

The Application layer is probably the most easily misunderstood layer of the model. This top layer defines the language and syntax that programs use to communicate with other programs. The application layer represents the purpose of communicating in the first place. For example, a program in a client workstation uses commands to request data from a program in the server. Common functions at this layer are opening, closing, reading and writing files, transferring files and e-mail messages, executing remote jobs and obtaining directory information about network resources etc.

Presentation layer (Layer 6)

The Presentation layer performs code conversion and data reformatting (syntax translation). It is the translator of the network; it makes sure the data is in the correct form for the receiving application.

When data are transmitted between different types of computer systems, the presentation layer negotiates and manages the way data are represented and encoded. For example, it provides a common denominator between ASCII and EBCDIC machines as well as between different floating point and binary formats. Sun's XDR and OSI's ASN.1 are two protocols used for this purpose. This layer is also used for encryption and decryption. It also provides security features through encryption and decryption.

Session layer (Layer 5)

The Session layer decides when to turn communication on and off between two computers. It provides the mechanism that controls the data-exchange process and coordinates the interaction (communication) between them in an orderly manner.

It sets up and clears communication channels between two communicating components. It determines one-way or two-way communications and manages the dialogue between both parties; for example, making sure that the previous request has been fulfilled before the next one is sent. It also marks significant parts of the transmitted data with checkpoints to allow for fast recovery in the event of a connection failure.

Transport layer (Layer 4)

The transport layer is responsible for overall end-to-end validity and integrity of the transmission i.e., it ensures that data is successfully sent and received between two computers. The lower data link layer (layer 2) is only responsible for delivering packets from one node to another. Thus, if a packet gets lost in a router somewhere in the enterprise Internet, the transport layer will detect that. It ensures that if a 12MB file is sent, the full 12MB is received.

If data is sent incorrectly, this layer has the responsibility of asking for retransmission of the data. Specifically, it provides a network-independent, reliable message-independent, reliable message-interchange service to the top three application-oriented layers. This layer acts as an interface between the bottom and top three layers. By providing the session layer (layer 5) with a reliable message transfer service, it hides the detailed operation of the underlying network from the session layer.

Network layer (Layer 3)

The network layer establishes the route between the sending and receiving stations. The unit of data at the network layer is called a packet. It provides network routing and flow and congestion functions across computer-network interface.

It makes a decision as to where to route the packet based on information and calculations from other routers, or according to static entries in the routing table. It examines network addresses in the data instead of physical addresses seen in the Data Link layer.

The Network layer establishes, maintains, and terminates logical and/or physical connections.

The network layer is responsible for translating logical addresses, or names, into physical addresses.

The main device found at the Network layer is a router.

Data link layer (Layer 2)

The data link layer groups the bits that we see on the Physical layer into Frames. It is primarily responsible for error-free delivery of data on a hop. The Data link layer is split into two sub-layers i.e., the Logical Link Control (LLC) and Media Access Control (MAC).

The Data-Link layer handles the physical transfer, framing (the assembly of data into a single unit or block), flow control and error-control functions (and retransmission in the event of an error) over a single transmission link; it is responsible for getting the data packaged and onto the network cable. The data link layer provides the network layer (layer 3) reliable information-transfer capabilities.

The main network device found at the data link layer is a bridge. This device works at a higher layer than the repeater and therefore is a more complex device. It has some understanding of the data it receives and can make a decision based on the frames it receives as to whether it needs to let the information pass, or can remove the information from the network. This means that the amount of traffic on the medium can be reduced and therefore, the usable bandwidth can be increased.

Physical layer (Layer 1)

The data units on this layer are called bits. This layer defines the mechanical and electrical definition of the network medium (cable) and network hardware. This includes how data is impressed onto the cable and retrieved from it.

The physical layer is responsible for passing bits onto and receiving them from the connecting medium. This layer gives the data-link layer (layer 2) its ability to transport a stream of serial data bits between two communicating systems; it conveys the bits that moves along the cable. It is responsible for ensuring that the raw bits get from one place to another, no matter what shape they are in, and deals with the mechanical and electrical characteristics of the cable.

This layer has no understanding of the meaning of the bits, but deals with the electrical and mechanical characteristics of the signals and signalling methods.

The main network device found the Physical layer is a **repeater**. The purpose of a repeater (as the name suggests) is simply to receive the digital signal, reform it, and retransmit the signal. This has the effect of increasing the maximum length of a network, which would not be possible due to signal deterioration if, a repeater were not available. The repeater, simply regenerates cleaner digital signal so it doesn't have to understand anything about the information it is transmitting, and processing on the repeater is non-existent.

An example of the Physical layer is RS-232.

Each layer, with the exception of the physical layer, adds information to the data as it travels from the Application layer down to the physical layer. This extra information is called a header. The physical layer does not append a header to information because it is concerned with sending and receiving information on the individual bit level.

We see that the data for each layer consists of the header and data of the next higher layer. Because the data format is different at each layer, different terms are commonly used to name the data package at each level. *Figure 10* summarises these terms layer by layer.

LAYER	**DATA PACKAGE NAME**
Application layer	Application Protocol Data Unit
Presentation layer	Presentation Protocol Data Unit. Generic name is Protocol Data Unit (PDU). For example, at session layer it is S-PDU.
Session layer	Session Protocol Data Unit
Transport layer	Segment
Network layer	Datagram, packet
Data link layer	Frame
Physical layer	Bit

Figure 10: Data package names in the OSI reference model.

OSI Protocols

The OSI model provides a conceptual framework for communication between computers, but the model itself is not a method of communication. Actual communication is made possible by using communication protocols. In the context of data networking, a *protocol* is a formal set of rules and conventions that governs how computers exchange information over a network medium. A protocol implements the functions of one or more of the OSI layers. A wide variety of communication protocols exist, but all tend to fall into one of the following groups: *LAN protocols*, *WAN protocols*, *network protocols*, and *routing protocols. LAN protocols* operate at the network and data link layers of the OSI model and define communication over the various LAN media. *WAN protocols* operate at the lowest three layers of the OSI model and define communication over the various wide-area media. *Routing protocols* are network-layer protocols that are responsible for path determination and traffic switching. Finally, *network protocols* are the various upper-layer protocols that exist in a given protocol suite.

Working of OSI Model

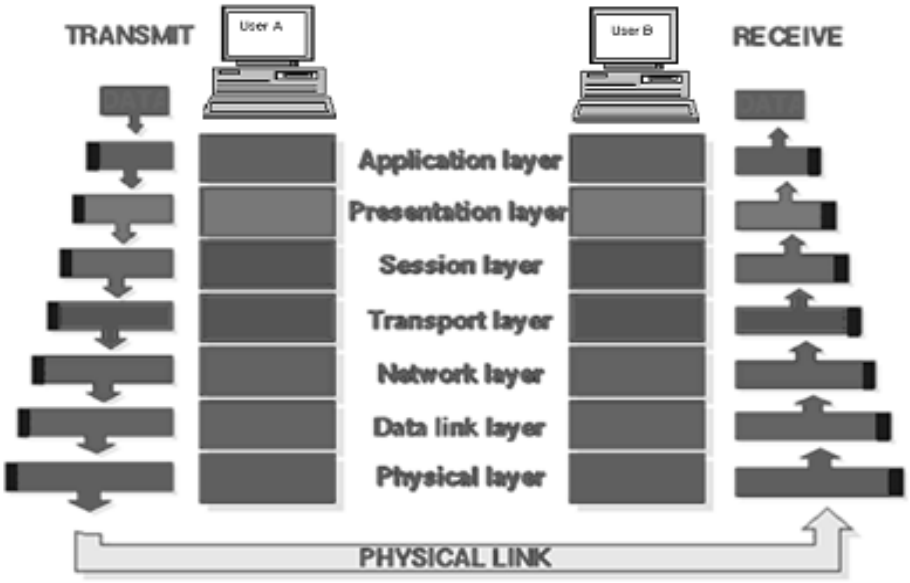

Figure 11: Working of OSI Reference Model

Information being transferred from a software application in one computer system to software application in another must pass through each of the OSI layers. Each layer communicates with three other OSI layers i.e., the layer directly above it, the layer directly below it, and its peer layer in other networked systems. If, for example, in *Figure 10*, a software application in System A has information to transmit to a software application in System B, the application program in System A will pass its information to the application layer (Layer 7) of System A. The application layer then passes the information to the presentation layer (Layer 6); the presentation layer reformats the data if required such that B can understand it. The formatted data is passed to the session layer (Layer 5), which in turn requests for connection establishment between session layers of A and B, it then passes the data to the transport layer. The transport layer breaks the data into smaller units called segments and sends them to the Network layer. The Network layer selects the route for transmission and if, required breaks the data packets further. These data packets are then sent to the Data link layer that is responsible for encapsulating the data packets into data frames. The Data link layer also adds source and destination addresses with error checks to each frame, for the hop.

The data frames are finally transmitted to the physical layer. In the physical layer, the data is in the form of a stream of bits and this is placed on the physical network medium and is sent across the medium to System B.

B receives the bits at its physical layer and passes them on to the Data link layer, which verifies that no error has occurred. The Network layer ensures that the route selected for transmission is reliable, and passes the data to the Transport layer. The function of the Transport layer is to reassemble the data packets into the file being transferred and then, pass it on to the session layer. The session layer confirms that the transfer is complete, and if so, the session is terminated.

The data is then passed to the Presentation layer, which may or may not reformat it to suit the environment of B and sends it to the Application layer. Finally the Application layer of System B passes the information to the recipient Application program to complete the communication process.

Interaction between different layers of OSI model

A given layer in the OSI layers generally communicates with three other OSI layers: the layer directly above it, the layer directly below it, and its peer layer in an other networked computer system. The data link layer in System A, for example, communicates with the network layer of System A, the physical layer of System A, and the data link layer in System B.

Services provided by OSI layers

One OSI layer communicates with another layer to make use of the services provided by the second layer. The services provided by adjacent layers help a given OSI layer communicate with its peer layer in other computer systems. Three basic elements are involved in layer services: the service user, the service provider, and the service access point (SAP).

In this context, the service *user* is the OSI layer that requests services from an adjacent OSI layer. The service *provider* is the OSI layer that provides services to service users. OSI layers can provide services to multiple service users. The *SAP* is a conceptual location at which one OSI layer can request the services of another OSI layer.

OSI Model Layers and Information Exchange

The seven OSI layers use various forms of control information to communicate with their peer layers in other computer systems. This *control information* consists of specific requests and instructions that are exchanged between peer OSI layers.

Control information typically takes one of two forms: headers and trailers. Headers are prepended to data that has been passed down from upper layers. Trailers are appended to data that has been passed down from upper layers.

Headers, trailers, and data are relative concepts, depending on the layer that analyses the information unit. At the network layer, an information unit, for example, consists of a Layer 3 header and data. At the data link layer, however, all the information passed down by the network layer (the Layer 3 header and the data) is treated as data. In other words, the data portion of an information unit at a given OSI layer potentially can contain headers, trailers, and data from all the higher layers. This is known as encapsulation.

1.7.2 TCP/IP Reference Model

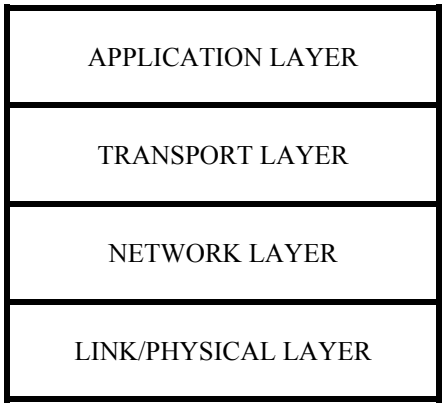

Figure 12: Layers of TCP/IP reference model

TCP/IP stands for Transmission Control Protocol / Internet Protocol. It is a protocol suite used by most communications software. TCP/IP is a robust and proven technology that was first tested in the early 1980s on ARPA Net, the U.S. military's Advanced Research Projects Agency network, and the world's first packet-switched network. TCP/IP was designed as an open protocol that would enable all types of computers to transmit data to each other via a common communications language.

TCP/IP is a layered protocol similar to the ones used in all the other major networking architectures, including IBM's SNA, Windows' NetBIOS, Apple's AppleTalk, Novell's NetWare and Digital's DECnet. The different layers of the TCP/IP reference model are shown in *Figure 13*. Layering means that after an application initiates the communications, the message (data) to be transmitted is passed through a number of stages or layers until it actually moves out onto the wire. The data are packaged with a different header at each layer. At the receiving end, the corresponding programs at each protocol layer unpack the data, moving it "back up the stack" to the receiving application.

TCP/IP is composed of two major parts: TCP (Transmission Control Protocol) at the transport layer and IP (Internet Protocol) at the network layer. TCP is a connection-oriented protocol that passes its data to IP, which is a connectionless one. TCP sets up a connection at both ends and guarantees reliable delivery of the full message sent. TCP tests for errors and requests retransmission if necessary, because IP does not.

An alternative protocol to TCP within the TCP/IP suite is UDP (User Datagram Protocol), which does not guarantee delivery. Like IP, it is also connectionless, but very useful for real-time voice and video, where it doesn't matter if a few packets get lost.

Layers of TCP/IP reference model

Application Layer (Layer 4)

The top layer of the protocol stack is the application layer. It refers to the programs that initiate communication in the first place. TCP/IP includes several application layer protocols for mail, file transfer, remote access, authentication and name resolution. These protocols are embodied in programs that operate at the top layer just as any custom-made or packaged client/server application would.

There are many Application Layer protocols and new protocols are always being developed.

The most widely known Application Layer protocols are those used for the exchange of user information, some of them are:

- **The HyperText Transfer Protocol (HTTP)** is used to transfer files that make up the Web pages of the World Wide Web.
- **The File Transfer Protocol (FTP)** is used for interactive file transfer.
- **The Simple Mail Transfer Protocol (SMTP)** is used for the transfer of mail messages and attachments.
- **Telnet,** is a terminal emulation protocol, and, is used for remote login to network hosts.

Other Application Layer protocols that help in the management of TCP/IP networks are:

- **The Domain Name System (DNS),** which, is used to resolve a host name to an IP address.
- **The Simple Network Management Protocol (SNMP)** which is used between network management consoles and network devices (routers, bridges, and intelligent hubs) to collect and exchange network management information.

Examples of Application Layer interfaces for TCP/IP applications are Windows Sockets and NetBIOS. Windows Sockets provides a standard application-programming interface (API) under the Microsoft Windows operating system. NetBIOS is an industry standard interface for accessing protocol services such as sessions, datagrams, and name resolution.

Transport Layer (Layer 3)

The Transport Layer (also known as the Host-to-Host Transport Layer) is responsible for providing the Application Layer with session and datagram communication services.

TCP/IP does not contain Presentation and Session layers, the services are performed if required, but they are not part of the formal TCP/IP stack. For example, Layer 6 (Presentation Layer) is where data conversion (ASCII to EBCDIC, floating point to binary, etc.) and encryption /decryption is performed. Layer 5 is the Session Layer, which is performed in layer 4 in TCP/IP. Thus, we jump from layer 7 of OSI down to layer 4 of TCP/IP.

From Application to Transport Layer, the application delivers its data to the communications system by passing a stream of data bytes to the transport layer along with the socket of the destination machine.

The core protocols of the Transport Layer are TCP and the User Datagram Protocol (UDP).

- **TCP:** TCP provides a one-to-one, connection-oriented, reliable communications service. TCP is responsible for the establishment of a TCP connection, the sequencing and acknowledgment of packets sent, and the recovery of packets lost during transmission.
- **UDP:** UDP provides a one-to-one or one-to-many, connectionless, unreliable communications service. UDP is used when the amount of data to be transferred is small (such as the data that would fit into a single packet), when the overhead of establishing a TCP connection is not desired, or when the applications or upper layer protocols provide reliable delivery.

The Transport Layer encompasses the responsibilities of the OSI Transport Layer and some of the responsibilities of the OSI Session Layer.

Internet Layer (Layer 2)

The Internet layer handles the transfer of information across multiple networks through the use of gateways and routers. The Internet layer corresponds to the part of the OSI network layer that is concerned with the transfer of packets between machines that are connected to different networks. It deals with the routing of packets across these networks as well as with the control of congestion. A key aspect of the Internet layer is the definition of globally unique addresses for machines that are attached to the Internet.

The Internet layer provides a single service namely, best-effort connectionless packet transfer. IP packets are exchanged between routers without a connection setup; the packets are routed independently and so they may traverse different paths. For this reason, IP packets are also called datagrams. The connectionless approach makes the system robust; that is, if failures occur in the network, the packets are routed around the points of failure; hence, there is no need to set up connections. The gateways that interconnect the intermediate networks may discard packets when congestion occurs. The responsibility for recovery from these losses is passed on to the Transport Layer.

The core protocols of the Internet Layer are IP, ARP, ICMP, and IGMP.

- **The Internet Protocol (IP)** is a routable protocol responsible for IP addressing and the fragmentation and reassembly of packets.
- **The Address Resolution Protocol (ARP)** is responsible for the resolution of the Internet Layer address to the Network Interface Layer address, such as a hardware address.
- **The Internet Control Message Protocol (ICMP)** is responsible for providing diagnostic functions and reporting errors or conditions regarding the delivery of IP packets.
- **The Internet Group Management Protocol (IGMP)** is responsible for the management of IP multicast groups.

The Internet Layer is analogous to the Network layer of the OSI model.

Link/Physical Layer (Layer 1)

The Link/Physical Layer (also called the Network Access Layer) is responsible for placing TCP/IP packets on the network medium and receiving TCP/IP packets of the network medium. TCP/IP was designed to be independent of the network access method, frame format, and medium. In this way, TCP/IP can be used to connect differing network types. This includes LAN technologies such as Ethernet or Token Ring and WAN technologies such as X.25 or Frame Relay. Independence from any specific network technology gives TCP/IP the ability to be adapted to new technologies such as Asynchronous Transfer Mode (ATM).

The Network Interface Layer encompasses the Data Link and Physical layers of the OSI Model. **Note,** that the Internet Layer does not take advantage of sequencing and acknowledgement services that may be present in the Data Link Layer. An unreliable Network Interface Layer is assumed, and reliable communications through session establishment and the sequencing and acknowledgement of packets is the responsibility of the Transport Layer.

Comparison between OSI and TCP/IP reference model

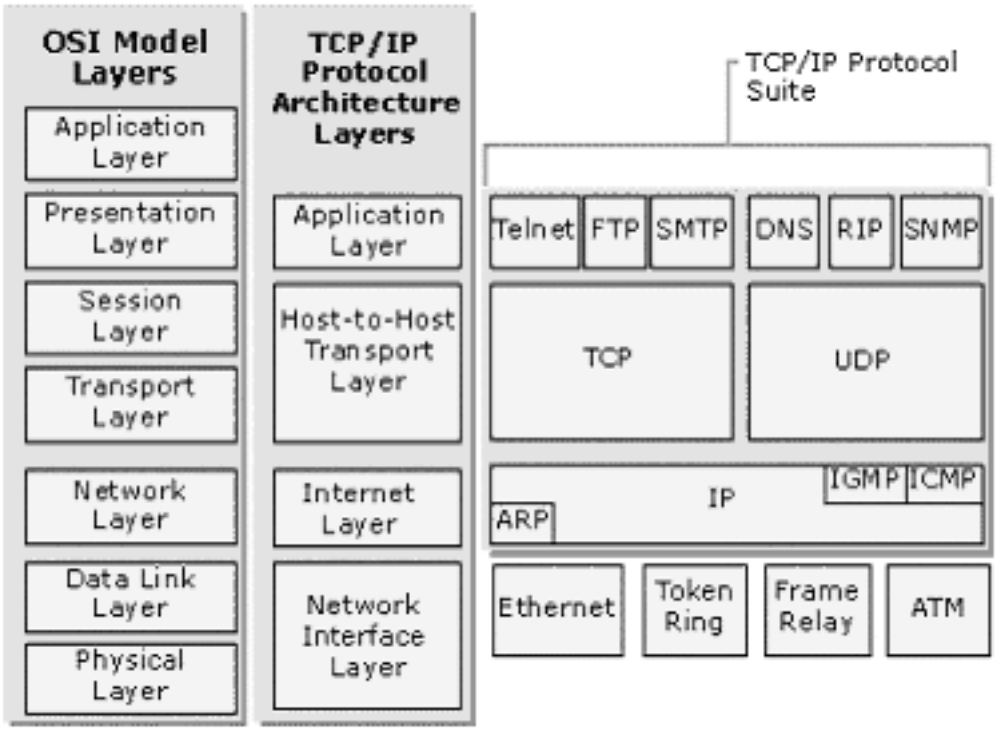

Figure 13: Comparison between OSI & TCP/IP reference model

Both OSI and TCP/IP reference models are based on the concept of a stack of protocols. The functionality of the layers is almost similar. In both models the layers are there to provide an end-to-end network-independent transport service to processes wishing to communicate with each other.

The Two models have many differences. An obvious difference between the two models is the number of layers: the OSI model has seven layers and the TCP/IP has four layers. Both have (inter) network, transport, and application layers, but the other layers are different. OSI uses strict layering, resulting in vertical layers whereas TCP/IP uses loose layering resulting in horizontal layers. The OSI model supports both connectionless and connection-oriented communication in the network layer, but only connection-oriented communication at the transport layer. The TCP/IP model has only one mode in network layer (connectionless), but supports both modes in the transport layer. With the TCP/IP model, replacing IP by a substantially different protocol would be virtually impossible, thus, defeating one of the main purposes of having layered protocols in the first place.

The OSI reference model was devised before the OSI protocols were designed. The OSI model was not biased toward one particular set of protocols, which made it quite general. The drawback of this ordering is that the designers did not have much experience with the subject, and did not have a good idea of the type of functionality to put in a layer. With TCP/IP the reverse was true: the protocols came first, and the model was really just a description of the existing protocols. There was no problem with the protocols fitting the model. The only drawback was that the model did not fit any other protocol stacks.

Figure 14 summarises the basic differences between OSI and TCP/IP reference models.

OSI MODEL	TCP/IP MODEL
Contains 7 Layers	Contains 4 Layers
Uses Strict Layering resulting in vertical layers.	Uses Loose Layering resulting in horizontal layers.
Supports both connectionless & connection-oriented communication in the Network layer, but only connection-oriented communication in Transport Layer	Supports only connectionless communication in the Network layer, but both connectionless & connection-oriented communication in Transport Layer
It distinguishes between Service, Interface and Protocol.	Does not clearly distinguish between Service, Interface and Protocol.
Protocols are better hidden and can be replaced relatively easily as technology changes (No transparency)	Protocols are not hidden and thus cannot be replaced easily. (Transparency) Replacing IP by a substantially different protocol would be virtually impossible
OSI reference model was devised before the corresponding protocols were designed.	The protocols came first and the model was a description of the existing protocols

Figure 14: Difference between OSI and TCP/IP reference model

Some of the drawbacks of OSI reference model are:

- All layers are not roughly, of equal size and complexity. In practise, the session layer and presentation layer are absent from many existing architectures.
- Some functions like addressing, flow control, retransmission are duplicated at each layer, resulting in deteriorated performance.
- The initial specification of the OSI model ignored the connectionless model, thus, leaving much of the LANs behind.

Some of the drawbacks of TCP/IP model are:

- TCP/IP model does not clearly distinguish between the concepts of service, interface, and protocol.
- TCP/IP model is not a general model and therefore it cannot be used to describe any protocol other than TCP/IP.
- TCP/IP model does not distinguish or even mention the Physical or the Data link layer. A proper model should include both these layers as separate.

☞ Check Your Progress 3

1) Give two reasons for using layered protocols.

……

……

……

2) Explain the OSI reference model in detail.

……

……

……

3) Explain the TCP/IP reference model in detail.

……

……

……

4) Bring out the differences between TCP and UDP.

……

……

……

1.8 NETWORK ARCHITECTURE

Depending on the architecture used Networks can be classified as Client/Server or Peer-to-Peer Networks.

1.8.1 Client/Server Architecture

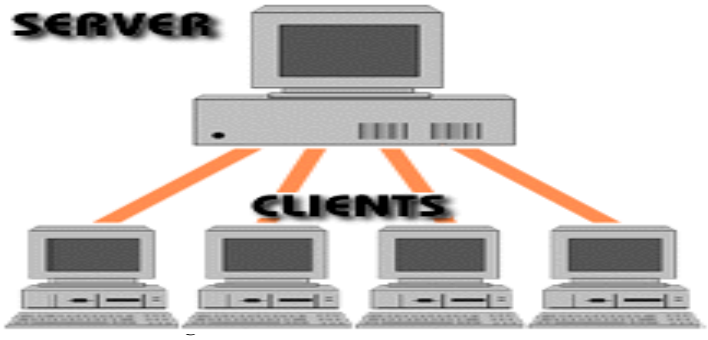

Figure 15: Client/Server architecture

Client/Server Architecture is one in which the client (personal computer or workstation) is the requesting machine and the server is the supplying machine, both of which are connected via a local area network (LAN) or wide area network (WAN). Since the early 1990s, client/server has been the buzzword for building applications on LANs in contrast to centralised minis and mainframes with dedicated terminals. A client/server network is called Centralised or Server based network. *Figure 15* shows the arrangement of computers in the client/server environment.

The client contains the user interface and may perform some or all of the application processing. Servers can be high-speed microcomputers, minicomputers or even mainframes. A database server maintains the databases and processes requests from the client to extract data from or update the database. An application server provides additional business processing for the clients.

The term client/server is sometimes used to contrast a peer-to-peer network, in which any client can also act as a server. In that case, a client/server entails having a dedicated server.

However, client/server architecture means more than dedicated servers. Simply downloading files from or sharing programs and databases on a server is not true client/server either. True client/server implies that the application was originally designed to run on a network and that the network infrastructure provides the same quality of service as traditional mini and mainframe information systems. *Figure 15* shows the arrangement of computers in the Client/Server system.

The network operating system (NOS) together with the database management system (DBMS) and transaction monitor (TP monitor) are responsible for integrity and security of these types of networks. Some of these products have gone through many client/server versions by now and have finally reached industrial strength.

Non-client/server

In non-client/server architecture, the server is nothing more than a remote disk drive. The user's machine does all the processing. If, many users routinely perform lengthy searches, this can bog down the network, because each client has to pass the entire database over the net. At 1,000 bytes per record, a 10,000 record database requires 10MB of data be transmitted.

Two-tier client/server

Two-tier client/server is really the foundation of client/server. The database processing is done in the server. An SQL request is generated in the client and transmitted to the server. The DBMS searches locally and returns only matching

records. If 50 records met the criteria, only 50K would be transmitted. This reduces traffic in the LAN.

Three-tier client/server

Many applications lend themselves to centralised processing. If, they contain proprietary algorithms, security is improved. Upgrading is also simpler. Sometimes, programs are just too demanding to be placed into every client PC. In three-tier client/server, application processing is performed in one or more servers.

1.8.2 Peer-to-Peer Architecture

Figure 16: Peer-to-peer architecture

A type of network in which each workstation has equal capabilities and responsibilities is called peer-to-peer network. *Figure 16* shows the arrangement of computers in a peer-to-peer environment. Here each workstation acts as both a client and a server. There is no central repository for information and there is no central server to maintain. Data and resources are distributed throughout the network, and each user is responsible for sharing data and resources connected to their system. This differs from client/server architectures, in which some computers are dedicated to serving the others. Peer-to-peer networks are generally simpler and less expensive, but they usually do not offer the same performance under heavy loads. A peer-to-peer network is also known as a Distributed network.

1.9 EXAMPLE NETWORKS

Nowdays, as computers are extensively used in almost every field, we have many different types of networks. Some of them are public networks, research networks, and co-operative networks, commercial or corporate networks. We can distinguish between different networks on the basis of their history, administration, facilities offered, technical design and the people who use them (user communities). Here we shall discuss some of the popular networks, such as, Novell NetWare, ARPANET, Internet, ATM network etc.

1.9.1 Novell Netware

Novell NetWare is the most popular network system in the PC world. Novell NetWare contains the protocols that are necessary to allow communication between different types of PC's and devices. There are several versions of NetWare. The earlier versions NetWare 286 version 2.X was written to run on 286 machines. NetWare 386 versions 3.X were written to run on 386 and 486 machines. The most recent version NetWare 4.X can probably run on almost any type of machine.

Novell Networks are based on the client/server model in which at least one computer functions as a network file server, which runs all of the NetWare protocols and maintains the networks shared data on one or more disk drives. File servers generally allow users on other PC's to access application software or data files i.e., it provides services to other network computers called clients.

There are two types of file servers:

- Dedicated file servers.
- Non-dedicated file servers.

Dedicated File Servers: Dedicated file server runs only NetWare and do not run any other software, such as Windows application. Dedicated file servers are mostly used in large networks, because, in large networks, one extra client is less significant and a dedicated server can handle a larger number of requests more efficiently. In large networks security is one of the major concerns and providing a clear distinction between client and server hardware provides greater security.

Non-dedicated File Server: Non-dedicated file server can run both applications and NetWare. It is useful in small networks because it allows the server to also act as a client and thus, increase the number of clients in the network by one.

There are many other servers within a Novell NetWare such as, Print server, Message server, Database server etc.

Print server: The job of the Print server is to allow users to access shared printers. A Print server manages both requests and printers.

Message server: The job of the Message server is to transfer email messages between a client's PC and a user's mailbox.

Database server: A database server manages database files i.e., it adds, deletes and modifies records in the database; queries the database and generates the result required by the client; and transmits the results back to the client.

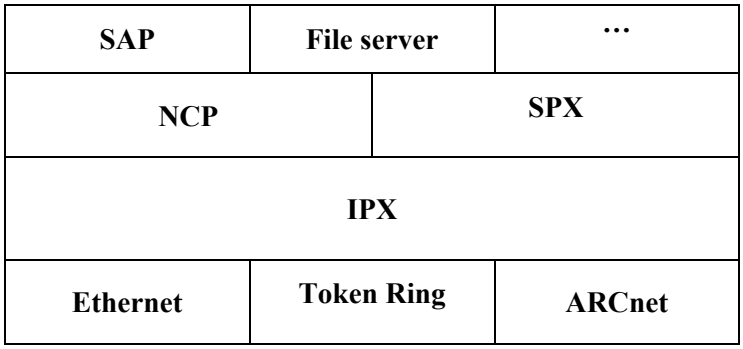

Figure 17: The Novell NetWare reference model

NetWare uses a proprietary protocol stack as shown in *Figure 17*. This model is based on the old Xerox Network System, XNS™ but with a lot of modifications.

The Physical and Data link layers can be chosen from various standards that are available, such as, the Ethernet, IBM Token ring, ARCnet.

The Network layer runs an unreliable connectionless internetwork protocol called Internet Packet eXchange (IPX). The IPX passes packets transparently from the

source to the destination, even if the source and destination are on different networks. The functioning of IPX is similar to IP, except that IPX uses12 byte addresses instead of 4 byte addresses.

The Transport layer contains a connection oriented transport protocol called NCP (Network Core Protocol). NCP is the heart of NetWare and provides various other services besides user data transport. It defines the type of requests that can be made and how the server responds to the request it eventually receives. The other protocol that is available in the transport layer is SPX (Sequenced Packet eXchange), which provides only transport. TCP is another option. Applications can choose anyone of them, for example, Lotus notes use SPX and File systems uses NCP.

Here, the Session and Presentation layers do not exist.

The Application layer contains various application protocols like SAP, File server etc.

.

1.9.2 ARPANET

ARPANET stands for Advanced Research Projects Agency (ARPA) Network. The network was developed in 1969 by ARPA and funded by the Department of Defence (DoD). In the mid-1960s at the height of the cold war, the DoD wanted a command and control network, which could survive the nuclear war. The traditional circuit switched telephone networks were considered too vulnerable, since the loss of one line would certainly terminate all conversations using them and might even partition the network.

The network (ARPANET) was chiefly experimental, and was used to research, develop and test networking technologies. The original network connected four host computers at four separate universities throughout the United States, enabling users to share resources and information. By 1972, there were 37 host computers connected to ARPANET. Also in that year, ARPA's name was changed to DARPA (Defence Advanced Research Projects Agency). In 1973, ARPANET went beyond the boundaries of the United States by making its first international connection to England and Norway. One goal of ARPANET was to devise a network that would still be operational, even if, part of the network failed. The research in this area resulted in a set of networking rules or protocols, called TCP/IP (Transmission Control Protocol/Internet Protocol).

TCP/IP is a set of protocols that govern how data is transmitted across networks. It also enables different types of computer operating systems such as DOS and UNIX to share data across a network.

ARPANET functioned as a "backbone" network allowing smaller local networks to connect to the backbone.

Once these smaller networks were connected to the backbone, they were in effect connected to each other.

In 1983, DARPA decided that TCP/IP would be the standard set of protocols used by computers connecting to ARPANET. This meant that any smaller networks (for example, a university network) that wanted to connect to ARPANET also had to use TCP/IP. TCP/IP was available for free and was increasingly used by networks. The spread of TCP/IP helped create the Internet, as we know it today, which is the network of networks that either use the TCP/IP protocols, or can interact with TCP/IP networks.

ARPANET continued to grow, encompassing many networks from universities and government institutions. To help manage this rapidly growing "network of networks", ARPANET was split into two networks in 1983:

- ARPANET continued to be a research and development network.
- MILNET an unclassified network reserved only for military sites. MILNET continues to serve this function.

In 1986, a faster backbone network called the NSFNET (National Science Foundation Network) was created. By 1989, there were over 10,000 host computers connected to the Internet.

Because of the success of the NSFNET, plans were made to phase out ARPANET. Many of the sites connected to ARPANET were absorbed by the NSFNET, and in 1990 ARPANET was officially dissolved.

1.9.3 Internet

When ARPANET and NSFNET were interconnected the number of networks, machines and users grew exponentially, many regional networks joined up and connections were made across many countries.

The internet is said to have been "officially" born around 1982 when the different networks (BITNET, EARN, etc.) agreed on using the TCP/IP protocol as a standard for their interconnections making it a network of networks and overcoming some of the previous cacophony of standards, protocols and increasing its coverage.

The word Internet was coined from the words "interconnection" and "network". Now Internet is the world's largest computer network. It is considered to be the network of networks, and is scattered all over the world. The computers connected to the Internet may communicate with each other using fiber optic cables, telephone lines, satellite links and other media.

The development of Internet is coordinated by a non-profit organisation called the Internet Society (ISOC). Its aim is to spread the use of Internet, keep statistics of its use, helpless developed countries in building their infrastructure and Internet-technology. The Internet Architecture Board (IAB), plans long term trends and keeps a record of the RFC (Request for Comments) documents on various technical solutions and protocols used in Internet. The development is also steered by the IETF (Internet Engineering Task Force), which has several sub-groups for handling various problems and planning new standards etc.

The rapid growth of Internet may also be due to several important factors:

1) Easy-to-use software - graphical browsers

2) Improved telecommunications connections

3) Rapid spread of automatic data processing, including electronic mail, bank ransfers, etc.

4) The Information Superhighway projects.

The Internet Society maintains a list of Internet service providers providing connections all over the world. There is one "universal" aspect of all computers connect to the Internet i.e., they all run the TCP/IP family of protocols.

The Internet Protocol (IP) gives the physical 32-bit address, which uniquely identifies an individual computer connected to the Internet, while Transmission Control Protocol (TCP) is a connection-oriented protocol, which takes care of the delivery and order of the packages. TCP also provides the port numbers for individual services within a computer.

The major information services provided by the Internet are (with the protocol in parentheses): electronic mail (SMTP), remote file copying (FTP), remote login, terminal connections (TELNET), menu-based file access (GOPHER), wide area information servers (WAIS, Z39.50), the World Wide Web (HTTP), and the Packet Internet Groper (PING).

There are three major ways to connect your computer to the Internet:

- dial up modem access to a computer connected to Internet,
- dial-up networking, and
- leased lines (usually from a local telephone company).

Switched Dial-Up Lines

The most common circuit provided by public communication carriers are dial-up telephone circuits. Subscribers send routing information i.e., the dialled number to the network, which connects them to the receiver, then follow this with the information (speech).

Switched circuits are not permanent. They exist only for the duration of the connection and are switched by the public network (it connects the circuits). Switched dial-up lines are not generally suited to data transmission, but are used heavily for some types of services (e.g., Bulletin Boards). Using a modem, a user can use their phone line to dial up a network provider via the phone line and connect to the Internet. At present speeds upto 56Kbps are possible over standard dial up telephone circuits.

Leased Lines

A leased line is a permanent non-switched end-to-end connection. Data is sent from one end to the other. It is not required to send routing information along with the data. Leased lines provide an instant guaranteed method of delivery. They are suited to high volume, high speed data requirements. The cost of the line (which is leased per month), is offset against that of toll or other rental charges. In addition, the leased line offers a significantly higher data transmission rate than the datel circuit.

Very high speeds can be achieved on leased lines. The cost varies, and goes up according to the capacity (speed in bits per second) that the customer requires.

Working of the Web

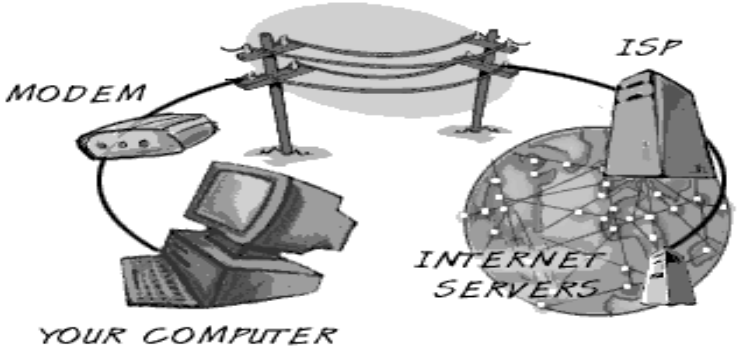

Figure 18: Working of internet

The Web physically consists of your personal computer, web browser software, a connection to an Internet service provider, computers called servers that host digital data and routers and switches to direct the flow of information.

The Web is known as a client-server system. Here, the Users computer is the client, and the remote computer that stores electronic files is the server.

The working of the Web can be explained by the following example:

Let's say you want to pay a visit to the IGNOU's website. First, you enter the address or URL of the website in your web browser (more about this in a while). Then your browser requests the web page from the web server. The IGNOU's server sends the data over the Internet to your computer. Your web browser interprets the data and displays it on your computer screen.

The glue that holds the Web together is called hypertext and hyperlinks. This feature allows electronic files on the Web to be linked so that you can easily jump between them. On the Web you can navigate through pages of information based on what interests you at that particular moment. This is commonly known as browsing or surfing the Net.

To access the Web you need software such as Netscape Navigator or Microsoft Internet Explorer. These are known as web browsers. Web pages are written in a computer language called HTML, which stands for Hypertext Markup Language.

1.9.4 ATM Network

Asynchronous Transfer Mode (ATM) is a network technology adopted by the telecommunication sector. It is a high-performance, cell-oriented switching and multiplexing technology that utilises fixed-length packets to carry different types of traffic. The data transfer takes place in the form of cells or packets of a fixed size (53 bytes). The cell used with ATM is relatively small compared to units used with older technologies. The small constant cell size allows ATM equipment to transmit video, audio, and computer data over the same network, and assures that no single type of data hogs the line.

ATM technology is used for both local and wide area networks (LANs and WANs) that support real-time voice and video as well as data. The topology uses switches that establish a logical circuit from end to end, which guarantees quality of service (QoS). However, unlike telephone switches that dedicate circuits end-to-end, unused bandwidth in ATM's logical circuits can be utilised when needed. For example, idle bandwidth in a videoconference circuit can be used to transfer data.

ATM is widely used as a backbone technology in carrier networks and large enterprises, but never became popular as a local network (LAN) topology. ATM is highly scalable and supports transmission speeds of 1.5, 25, 100, 155, 622, 2488 and 9953 Mbps. ATM is also running as slow as 9.6 Kbps between ships at sea. An ATM switch can be added into the middle of a switch fabric to enhance total capacity, and the new switch is automatically updated using ATM's PNNI routing protocol.

One of the important features of ATM is its ability to specify Quality of Service (QoS), allowing video and voice to be transmitted smoothly. It provides the following levels of service:

- Constant Bit Rate (CBR) guarantees bandwidth for realtime voice and video.
- Realtime variable Bit Rate (rt-VBR) supports interactive multimedia that requires minimal delays.
- Non-realtime variable bit rate (nrt-VBR) is used for bursty transaction traffic.

- Available Bit Rate (ABR) adjusts bandwidth according to congestion levels for LAN traffic.
- Unspecified Bit Rate (UBR) provides the best effort for non-critical data such as file transfers.

Advantages of ATM

- Flexible bandwidth allocation.
- Simple routing due to connection oriented technology.
- High bandwidth utilisation due to statistical multiplexing.
- Potential QOS (Quality Of Service) guarantees.

Disadvantages of ATM

- Overhead of cell header (5 bytes per cell).
- Complex mechanisms for achieving Quality of Service.
- Congestion may cause cell losses.
- It is costly compared to IP.

1. 10 TYPES OF COMPUTER NETWORKS

Computer Networks are mostly classified on the basis of the geographical area that the network covers, the topology used, the transmission media used and the computing model used.

Based on the geographical area covered the networks may be LAN, MAN, WAN.

1.10.1 Metropolitan Area Network (MAN)

Figure 19: Metropolitan area network

MAN (Metropolitan Area Network): Metropolitan Area Network is a Computer network designed for a town or city as shown in *Figure 19*. In terms of geographic area MAN's are larger than local-area networks (LANs), but smaller than wide-area networks (WANs). MAN's are usually characterised by very high-speed connections using fiber optical cable or other digital media.

The Typical Characteristics of a MAN are:

- Confined to a larger area than a LAN and can range from 10km to a few 100km in length.
- Slower than a LAN but faster than a WAN.

- Operates at a speed of 1.5 to 150 Mbps.
- Expensive equipment.
- Moderate error rates.

1.10.2 Wide Area Network (WAN)

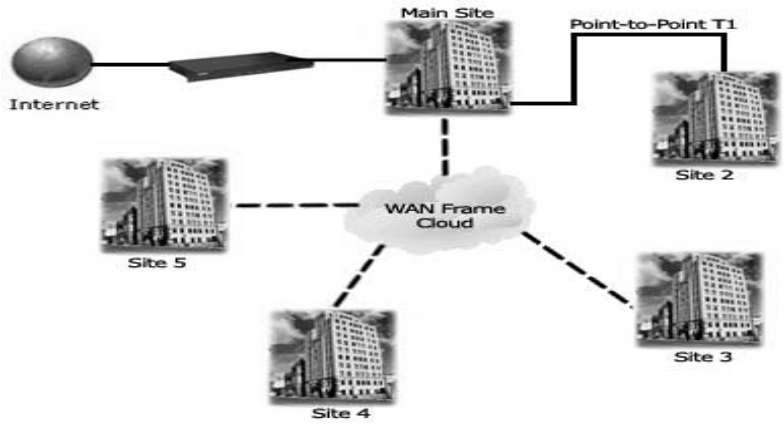

Figure 20: Wide area network

WAN (Wide Area Network): Wide Area Network is a computer network that spans a relatively large geographical area. Typically, a WAN consists of two or more local-area networks (LANs), which are depicted, in *Figure 20.* They can connect networks across cities, states or even countries.

Computers connected to a wide-area network are often connected through public networks, such as the telephone system. They can also be connected through leased lines or satellites.

The Typical characteristics of a WAN are:

- A WAN can range from 100km to 1000km and the speed between cities can vary form1.5 Mbps to 2.4 Gbps.
- WAN supports large number of computers and multiple host machines.
- Various segments of network are interconnected using sophisticated support devices like routers and gateways.
- Usually the speed is much slower than LAN speed.
- Highest possible error rate compared to LAN & MAN.

1.10.3 Comparison Between LAN, MAN, WAN and GAN

NETWORK	SIZE	TRANSMISSION MEDIA	MAXIMUM DISTANCE
Local Area Network	Confined to building or campus	Cable used	Covers up to 10 km
Metropolitan Area Network	Network confined to city or town	Different hardware & transmission media are used	Covers the area of a city or town
Wide Area Network	Larger than MAN	Telephone lines, radio waves, leased lines or satellites.	Covers a number of cities or countries

Figure 21: Comparison between different types of networks.

1.11 ADVANTAGES OF NETWORKS

Computers in a networked environment provide numerous advantages when compared to computers in a stand alone environment. The immense benefits that the computer networks provide are in the form of excellent sharing of computational resources, computational load, increased level of reliability, economy and efficient person-to-person communication.

Following are some of the major advantages of using computer networks.

Resource Sharing: The main aim of a computer network is to make all programs, equipment, and data available to anyone on the network without regard to the physical location of the resource and the user. Users need to share resources other than files, as well. A common example being printers. Printers are utilised only a small percentage of the time; therefore, companies don't want to invest in a printer for each computer. Networks can be used in this situation to allow all the users to have access to any of the available printers.

High Reliability: Computer networks provide high reliability by having alternative sources of supply. For example, all files could be replicated on two or three machines, so, if one of them is unavailable (due to hardware failure), the other copies could be used. In addition, the presence of multiple CPUs means that if one goes down, the others may be able to take over its work, although at reduced performance. For military, banking, air traffic control, nuclear reactor safety, and many other applications, the ability to continue operating in the face of hardware problems is of utmost importance.

Saving Money: Small computers have a much better price/performance ratio than larger ones. Mainframes are roughly a factor of ten faster than personal computers but they cost much more. This imbalance has caused many systems designers to build systems consisting of personal computers, one per user, with data kept on one or more shared **file server** machines. In this model, the users are called **clients**, and the whole arrangement is called the **client-server model.**

Scalability: The ability to increase the system performance gradually as the workload grows just by adding more processors. With centralised mainframes, when a system is full, it must be replaced by a larger one, usually at great expense and even greater disruption to the users. With client-server model, new clients and new servers can be added when needed.

Communication Medium: A computer network can provide a powerful communication medium among widely separated users. Using a computer network it is easy for two or more people who are working on the same project and who live far apart to write a report together. When one worker, makes a change to an on-line document, the others can see the change immediately, instead of waiting several days for a letter. Such a speedup makes cooperation among far-flung groups of people easy whereas previously it was impossible.

Increased Productivity: Networks increase productivity as several people can enter data at the same time, but they can also evaluate and process the shared data. So, one person can handle accounts receivable, and someone else processes the profit-and-loss statements.

☞ Check Your Progress 4

1) Briefly describe about NCP and IPX of Novell NetWare reference model.

……

……

……

……

2) List the basic components (equipments) in order to connect a computer to the Internet.

……

……

……

……

3) What are advantages of having small fixed size cells in ATM?

……

……

……

……

1.12 SUMMARY

In this unit we have learnt about the basic concepts of Networking. Here we discussed the different types of networks and the difference between them. Computer networks are basically classified as LAN, MAN, WAN depending on the geographical distance covered and depending on the various ways of interconnecting computers in a network (network topology) like Star, Bus, Ring, Tree, Mesh and cellular topologies.

We have seen the immense benefits that the computer networks provide in the form of excellent sharing of computational resources, computational load, increased level of reliability, economy and efficient person-to-person communication. Here we have briefly explained some of the network protocols which define a common set of rules and signals that computers on the network use to communicate with each other.

Standard network architecture is required for meaningful communication between end systems. We have discussed the two most widely used reference models i.e., the OSI reference model and the TCP/IP reference model. Nowadays, we come across different types of networks like Public networks, Research networks, Co-operative networks, Commercial networks etc. and we have learnt about some of the popular networks such as Novell NetWare, ARPANET, Internet, ATM network. Towards the end of this unit the concept of delays was also introduced.

1.13 SOLUTIONS/ANSWERS

Check Your Progress 1

1) Difference between Client/Server and Peer-to-Peer architecture

Client/Server	Peer-to-Peer
1. Also known as Centralised or Server based network.	1. Also known as distributed network.
2. Some computers in the network are dedicated to a particular task and are called Servers.	2. Each workstation has equal rights and responsibilities i.e., each work station acts as both a client and a server.
3. Only some computers are dedicated to serve others.	3. Data and resources are distributed throughout the network and each user is responsible for sharing data and resources connected to their system.
4. These networks are more expensive when compared to peer-to peer networks.	4. These networks are generally simpler and less expensive.

2) Some of the important aspects that should be considered while designing a network are:

- Cost reduction.
- High reliability.
- Increased productivity.
- Use of standard protocols etc.

3) Refer to Section 1.6

4) Difference between Broadcast and Point-to-Point networks.

Broadcast Networks: Broadcast networks have a single communication channel that is shared by all the machines on the network. In this type of network, short messages called packets, sent by any machine are received by all the machines on the network. The packet contains an address field, which specifies for whom the packet is intended. All the machines upon receiving a packet check for the address field, (if the packet is intended for itself), it processes it and if not the packet is just ignored.

Point-to-point or Switched networks: Point-to-point or Switched networks are those in which there are many connections between individual pairs of machines. In these networks when a packet travels from source to destination it may have to first visit one or more intermediate machines. Routing algorithms play an important role in Point-to-Point or Switched networks because often multiple routes of different lengths are available. Large networks are usually Point-to-Point networks.

Check Your Progress 2

1) Students are advised to write as an excercise.

2) The various types of networks are:

- **Local Area Network (LAN):** The network that spans a relatively small area, that is, in the single building or campus.
- **Metropolitan Area Network (MAN):** The type of computer network that is designed for a city or town.
- **Wide Area Network (WAN) :** A network that covers a large geographical area and covers different cities, states and sometimes even countries.

3) Comparison between Star and Tree Topology.

Star Topology	Tree Topology
1. All devices are connected directly to the central hub. 2. There is only one central hub. 3. In order for the nodes to communicate across the network data always passes through the central hub. 4. More cabling is required when compared to tree topology.	1. All devices are not connected to the central hub directly. 2. It is not always necessary for the data to pass through the central hub. 3. The addition of a secondary hub allows more devices to be attached to the central hub. 4. Less cabling required when compared to Star topology.

4) The best suited topology is Star Topology, reason out why.

Check Your Progress 3

1) The following are the reasons why layered protocols are useful.

- They help to reduce the design complexity of computer network.
- Help in understanding different concepts.

2) Refer to Section 7.1

3) Refer to Section 7.2

4) Difference between TCP and UDP

Transmission Control Protocol (TCP):

- Transmission Control Protocol is a highly reliable, connection oriented, end-to-end transport layer protocol.
- TCP provides message fragmentation and reassembly, and can accept messages of any length from upper layer protocols.

- TCP can maintain multiple conversations with upper layer protocols and can improve use of network bandwidth by combining multiple messages into the same segment.

User Datagram Protocol (UDP):

- User Datagram Protocol is a unreliable connectionless transport (host-to-host) layer protocol.
- UDP does not provide message acknowledgments, it simply transports datagrams.
- UDP is used when the amount of data to be transferred is small (such as the data that would fit into a single packet), when the overhead of establishing a TCP connection is not desired, or when the applications or upper layer protocols provide reliable delivery.
- Like TCP, UDP utilises port addresses to deliver datagrams.

Check Your Progress 4

1) Novell NetWare is best suited in a computer network consisting of PC's and it is based on the Client/Server model.
NCP (NetWare Core Protocol) is a connection oriented transport protocol. This is one of the important protocols of Novell NetWare. NCP defines the type of requests that can be made and how the server responds to these requests.

IPX (Internetwork Packet Exchange). This protocol is used to establish and maintain connections between network devices. IPX determines source and destination addresses and stores them in a packet along with the request. IPX does not guarantee packet delivery.

2) The basic equipments that are required in order to connect to the Internet is

- Computer
- Modem
- Connection to Internet provided by the Internet Service Provider.

3) One of the key aspects of ATM is that it uses 53 bytes fixed size cells. In which 48 bytes are for data and 5 bytes for header.
The advantages of using small fixed size cells are:

- It is easy for programmers to write programs that manipulate fixed-size record structures than writing ones, which deal with variable-size structures.
- Because of small cells, the queuing delay for high priority cell is reduced, because small cells do not occupy an outgoing link for lengthy periods.
- If the cell size is fixed the bytes arrive at a more consistent rate at the final destination as opposed to a bustier arrival pattern.
- The fixed cell size ensures that time-critical information such as voice or video is not adversely affected by long data frames or packets.
- The header is organised for efficient switching in high speed hardware implementations and carries payload type information, virtual circuit identifiers and header error check.

1.14 FURTHER READINGS

1) *Computer Networks*, Andrew S. Tenenbaum, PHI, New Delhi.
2) *Data and Computer Communication*, William Stalling, PHI, New Delhi.