

गोपनीय

गोपनीय

स्वावलम्बन

उत्तर प्रदेश राजर्षि टाउन मुक्त विश्वविद्यालय



HARDWARE TECHNOLOGY

BLISS-03
Computer Network and
Security Maintenance

UPRTOU

First Block : Network Basics
Second Block : Transmission and Network Elements
Third Block : Internet Connectivity
Fourth Block : Installation and Administration

Fifth Block : Software Solution of Peripherals
Sixth Block : Installation and Maintenance of I/O Devices
Seventh Block : Maintenance and Trouble-Shooting Of Monitor
Eighth Block : Maintenance and Trouble Shooting Of Printers

शान्तिपुरम (सेक्टर-एफ), फाफामऊ, इलाहाबाद- 211013



BLOCK

1

Network Basics

Unit 1

Introduction to Networking, Need, Advantages, and Types 5

Unit 2

Network Topologies 18

Unit 3

Network Protocols, Hardware and Software 29

Unit 4

Network Design and Configuration 50

Block Introduction:

This is the first block of the Data Communication and Networking. In this block networking concepts, topologies, protocols, hardware and software are introduced.

In unit 1, there is introduction to networking, need, advantages and their types on different basis.

In unit 2, different network topologies are presented with their comparison on the basis of advantages and disadvantages.

Unit 3 is mainly about network protocols, hardware and software. This includes defining network protocols and standards and describing network hardware and software.

Unit 4 is mainly focused about network design and configuration.

Unit I – Introduction to Networking, Need, Advantages and Types

What You Will Learn (Objectives)

After completing this unit you will be able to:

- Define LAN, Workgroup LAN, Departmental LAN, WAN, MAN, and Enterprise Network.
- Graphically represent LANs, MANs, and WANs.
- Describe the differences between a peer-to-peer network and a client/server network.

1.1 INTRODUCTION TO NETWORKS

Networks satisfy a broad range of purposes and meet various requirements. Some of the common objectives of computer communication networks are:

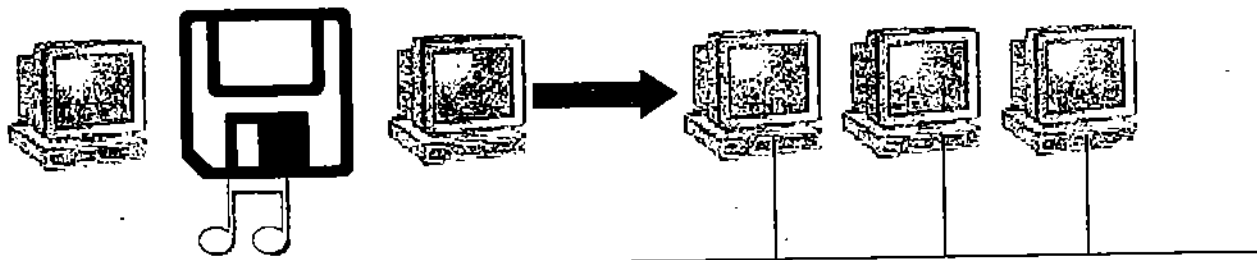
1. To provide sharing of geographically distant resources such as information, databases or processors (CPUs). Resources' sharing is perhaps the most common objective for providing networks, within the constraints of cost and reliability of transmission links.
2. To provide communication among users. Network users geographically far apart from each other can converse in an interactive session or send messages to each other.
3. To increase the reliability of processing capacity through backup and redundancy. If one processing unit breaks down, another physically distant processor (which is a backup of this unit) can take over.
4. To provide distributed processing capability, which means taking the processing out of a single large computer and distributing it to the place where the data is generated or where most of the operations are done. It is cost-effective as it eliminates most of the expensive large processors, and also saves on transmission cost.
5. To provide centralized management and allocation of resources.
6. To enable modular enhancement of computing resources. We can at any time add (or remove) an extra smaller and inexpensive computer and connect it to the network to increase the total computing capacity of the network. Doing the same thing in a big central computer is difficult and expensive.
7. Superior price/performance ratio. It is important to note here that there are still some applications which require a tremendous amount of processing capacity and are best handled by a powerful central computer and not by a large number of distributed small computers. Such jobs can be obtained over the network.

1.2 Basic Network Types

At a Glance

In the past, sharing files between computers occurred via *sneakernet*, literally walking a file stored on a floppy disk from one computer to another. In an effort to improve the transfer of information, the modern computer networks were born. A network is simply two or more computers connected together to share information and resources.

The Evolution of Modern Networking



SneakerNet

- Computer not Connected
- Data transferred by foot

Network

- Computer Connected
 - Data transferred electronically
-

Computer networks allow people to:

- Exchange information (for example, database, documents, and graphics) via connected workstations.
- Share resources equipment (for example, computers, printers and scanners).
- Use shared applications (for example, spreadsheets and word processing programs).
- Collaborate and communicate electronically.

There are many types of networks. Choosing the best network structure depends on an organization's productivity needs, its budgetary restrictions, and the types of resources to be connected. To make the best choices requires understanding the basic types of networks and their functions.

Terminology

- **Network** – Consists of physical media, routers and switches, which transfers data between a user and their resources.
- **Client** – A computer that uses the shared resources located on a network. The computer it accesses to obtain these shared resources is sometimes called a host or a server.
- **Client/Server Network** – In a client/server relationship, client computers request resources and information from a central, usually more powerful, server computer. The main alternative to peer-to-peer.
- **Enterprise Network** – A combination of LAN, Man, or WAN networks and nodes, managed and owned by a private organization.
- **Local Area Network** – Two or more computers linked together for the purpose of sharing information and resources in limited geographical coverage area/distance.
- **Metropolitan Area Networks** – A series of LAN networks connected together over a local geographical area.
- **Node** – A device connected to the network. It might be a printer, a workstation, a file server, a client, a mainframe, and so on.
- **Peer-to-Peer Network** – Computers that share resources and information equally. There are no powerful central computers (servers) sharing information and resources. The main alternative to client/server.
- **Resources** - Computer hardware (printers, computers, servers, and so on) and software applications shared over a network.
- **Server** – A computer on a network that shares resources with other computers. There are different types of servers such as file servers, database servers, print servers, and so on.
- **Wide Area Networks** – Two or more LANs or MANs connected together generally spanning a geographical area; they may even span the globe.
- **Workstation** – A standalone computer that has its own CPU, used for common computer practices. Your home or school computer is most likely a workstation. Workstations can be networked together.

Basic Network Types

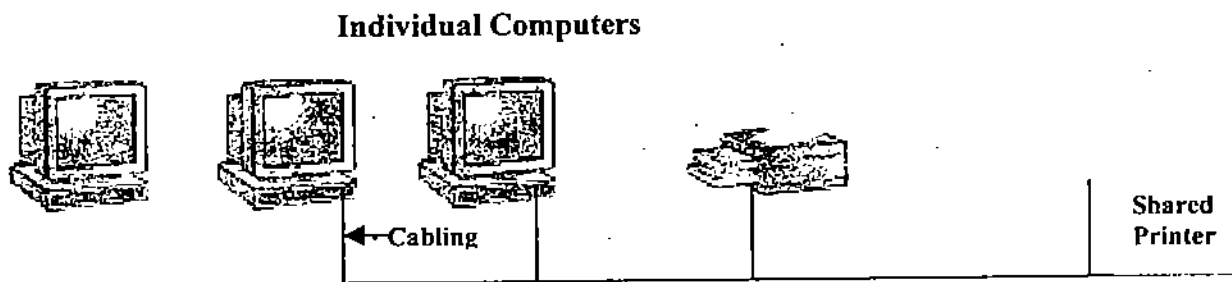
Schools, businesses, and other organizations select network types according to their individual and finances. Although network designs are quite diverse, basic network types are relatively few. They include the following:

- Local Area Networks
- Metropolitan Area Networks
- Wide Area Networks

1.2.1 Local Area Networks

A Local Area Networks (LAN) connects computers and hardware devices together over a small geographic area. These computers and hardware devices are frequently referred to as *nodes*.

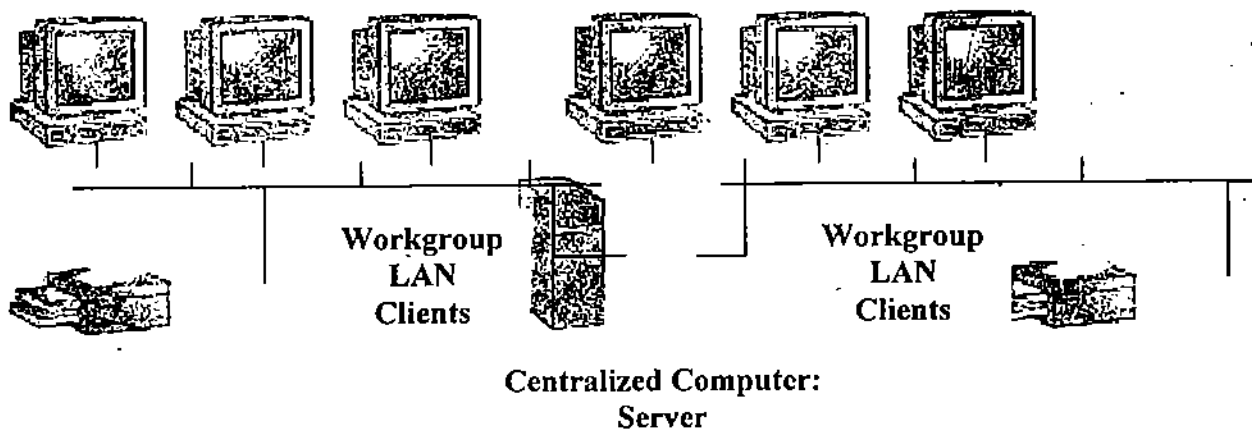
LAN: Peer-to-Peer Network



A LAN typically consists of several computers connected to one another, usually located in close proximity such as a computer lab. LANs are the most common form of networks found in most small businesses and schools. A LAN is often created for several computers to share an intermittently used resource such as a laser printer.

In earlier days of networking, most LANs used peer-to-peer networking to facilitate communication. A peer-to-peer network is built without the use of a central server, a computer used as the central storage unit for data. Each peer computer can share its resources with the other computers on the network independently. Today LANs often include central server computers.

LAN: Client/Server Network



In more complex networks client/server networking is usually employed. A centralized computer is used to facilitate the sharing of data and hardware with other computers on the network. These central computers are called *file servers*. The computers using the file server are called *clients*.

Check Your Understanding

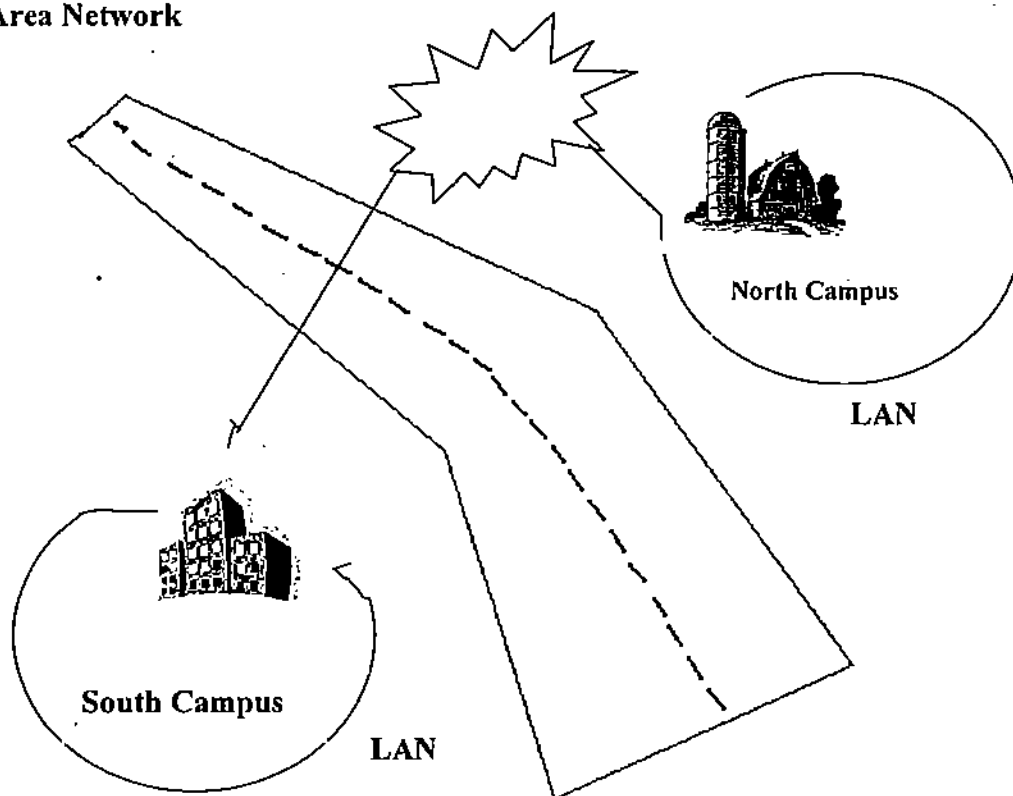
- Why have peer-to-peer networks been replaced overall by client/server networks?
- Describe in your own words, a Local Area Network.
- Sketch a Local Area network.

1.2.2 Metropolitan Area Networks

Two or more LANs in the same metropolitan area connected together are called Metropolitan Area Networks, MANs.

Using high-speed (1.544 Mbps to 155 Mbps) transmission, MANs connect smaller networks within a campus, a city, and/or neighboring suburbs.

Metropolitan Area Network



Companies usually pay an outside local carrier to supply the physical media necessary for transmitting data. The equipment and services provided by these vendors are usually on a monthly fee-for-service basis, with a one-time installation and set-up charge. One example is when a company leases telephone lines from a telecommunications company.

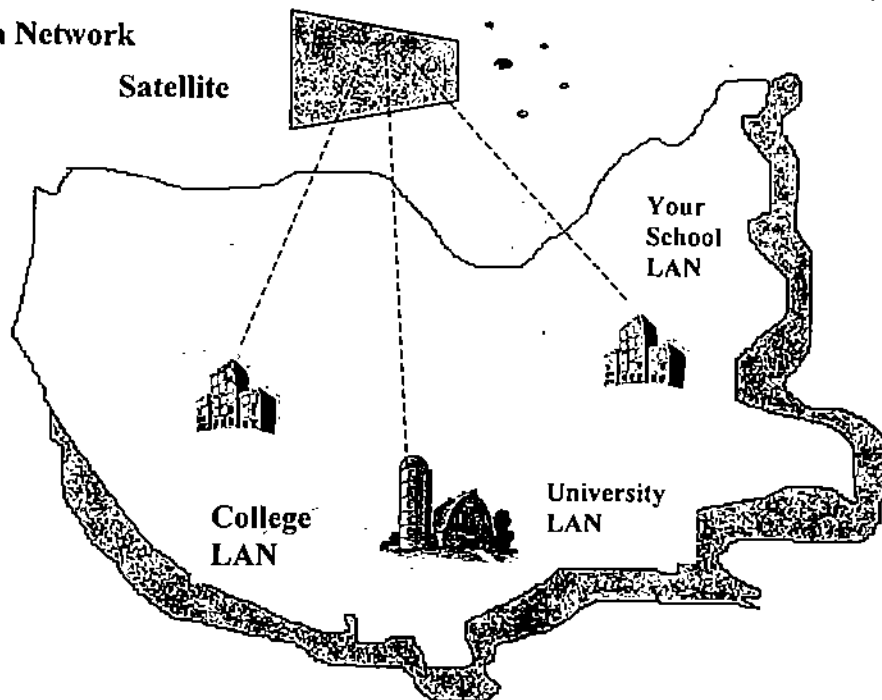
1.2.3 Wide Area Networks

Wide Area Networks (WANs) span large geographical distances, even over oceans or across continents. WANs overcome the distance limitations imposed by LANs.

WANs can be used to connect:

- LANs within school campus
- Networks within a city, a state, a country, or across the world.

Wide Area Network



WANs can use either analog (telephone lines) or digital (such as satellite transmission) signals, or a combination of both. WANs can be privately owned by large corporations or they can be public. One difference between public Mans and WANs is that the telephone company uses a long distance rather than local carrier.

1.2.4 Enterprise Networks

An Enterprise Network is the sum of the networked parts of an organization, encompassing all of the organization's LANs, MANs (Metropolitan Area Network) and WANs as well as clients, servers, printers, and other networked nodes.

Check Your Understanding

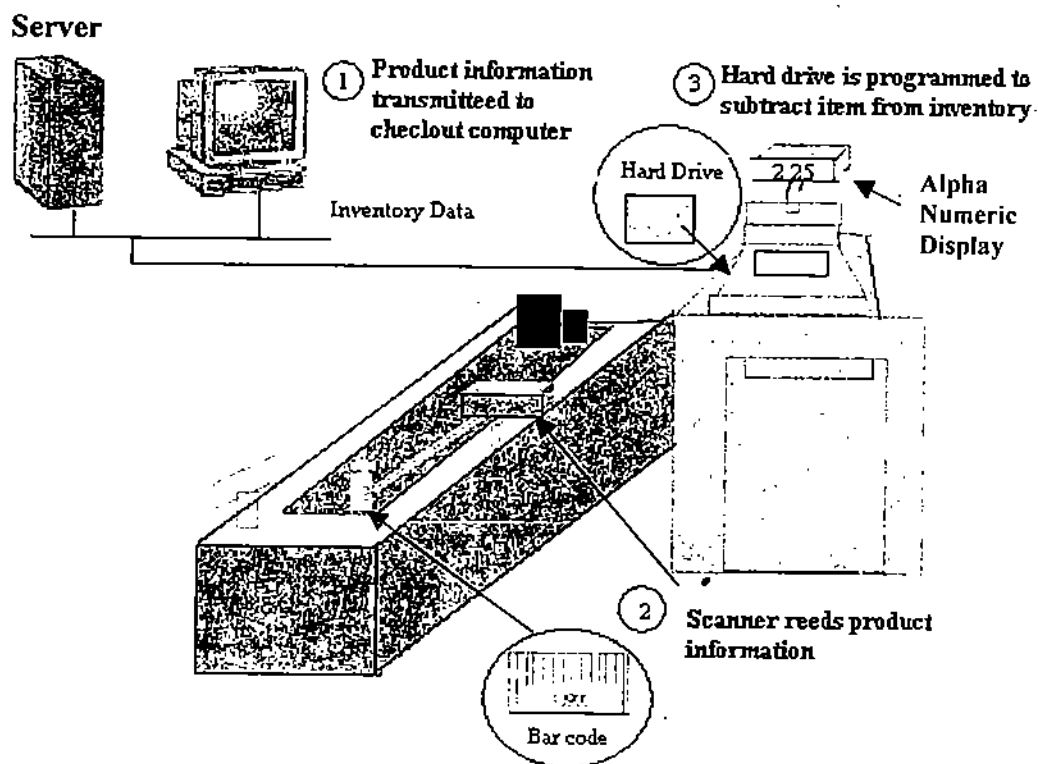
- Networks lower costs of doing work by allowing users to share resources. Sharing resources also increases efficiency. Give several examples of common resources a LAN might share.
- What security issues do you think might arise within a network environment?
- Which type of network might use peer-to-peer networking?
- Describe a situation where you might want to share resources with another individual.
- What is the difference between a WAN and Enterprise Network?
- Name at least two software protocols used for peer-to-peer networking.
- Diagram and label a Metropolitan area Network.

1.3 Networks at Work

When making a purchase at a modern supermarket chain, information can be exchanged between Local, Metropolitan, and Wide Area networks. Consider a typical supermarket transaction:

1. Each day, information, including item names, prices and stock numbers, is transmitted from the in-store server to each client checkout computer's hard drive.
2. As an item is slid across the scanner, a laser beam reads the product's bar code and display the name and price of the item on the alphanumeric display. The bar code price of the name and cost of the item, the name of the manufacturer, and so on.

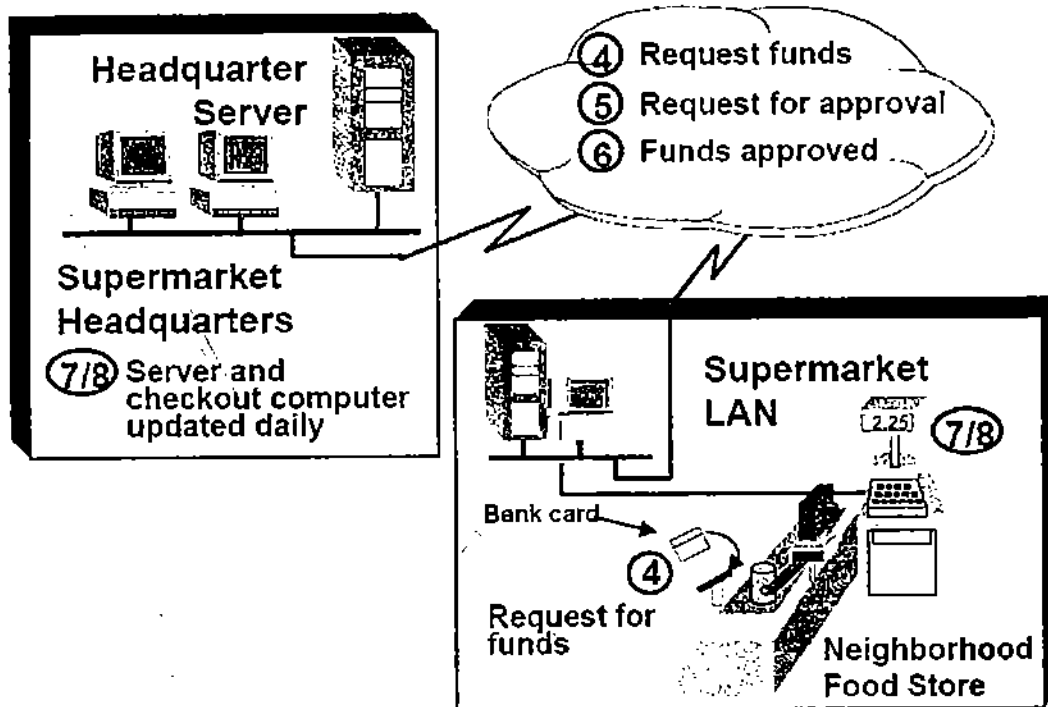
1.3.1 Supermarket Checkout Counter: Workgroup LAN



1. At the same time, the item purchased is subtracted from inventory.
2. To pay for the purchase, a debit or charge card may be slid through the swipe-card reader.
3. Approval is requested via the company's network from the central computer (server) at the headquarters. In this example, the headquarters is located in the same metropolitan area.

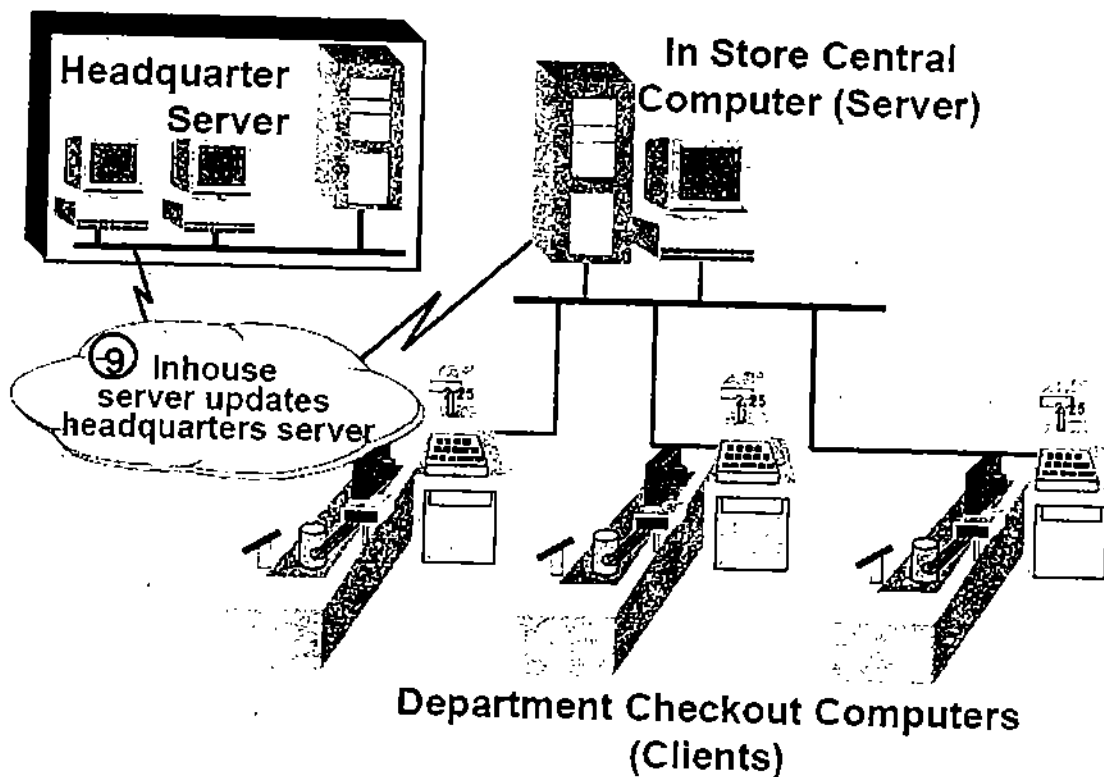
1.3.2 Supermarket's MAN

Supermarket's MAN



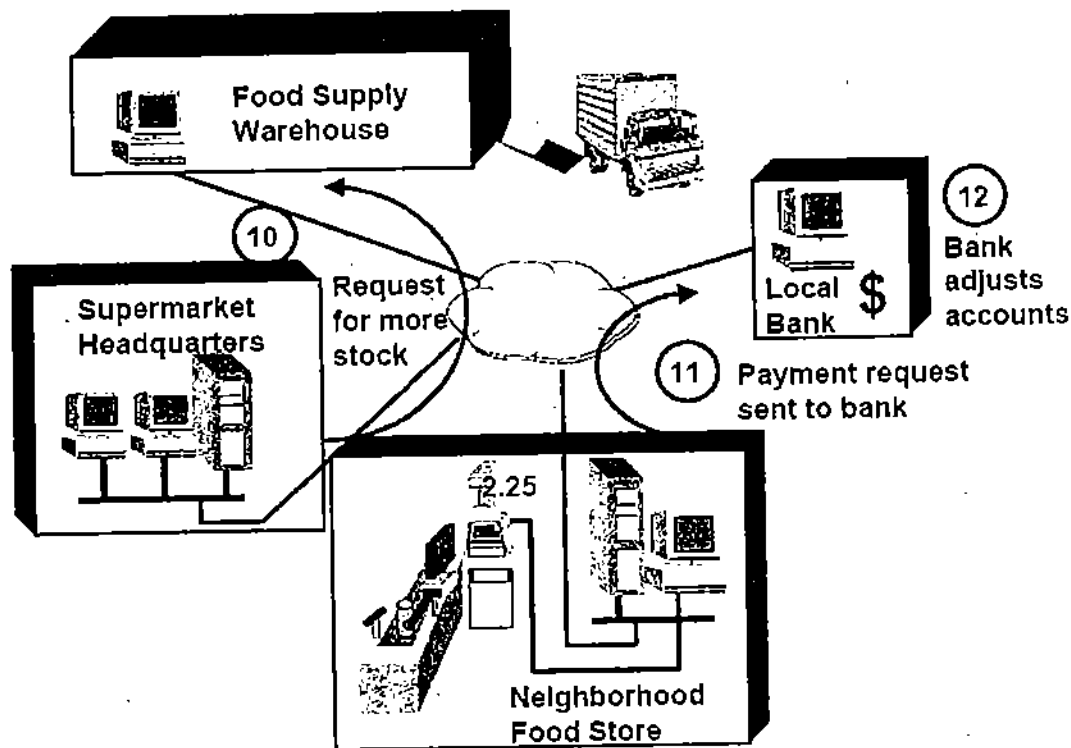
1. At the end of the business day, the in-store central computer (server) receives updated information from each cash register computer's hard drive.
2. New items, prices, and stock numbers are entered and transmitted to each computer for the next business day.
3. Data about the day's transactions are transmitted to the headquarters' server.

1.3.3 Supermarket's Departmental LAN



1. The headquarters' server, when stock is low, sends instructions to the warehouses for more supply deliveries.
2. Banking is also handled through the network. A payment request is sent electronically by the supermarket computer to your bank.

1.3.4 Supermarket's WAN



In turn, the bank subtracts the amount from your account and directs a credit to the supermarket's account.

Check Your Understanding

- Which type of network relationship is taking place between the checkout counter computer and the in-store central computer?
- Which type of relationship takes place among the swipe-card display computer, bar code reader computer, and checkout counter computer?
- Which type of network is used within the supermarket?
- Which type of network is used among the individual stores and the headquarters?
- Which type of network is used between the local banks, individual stores, the headquarters and the warehouse?

1.4 Summary

In this unit basic concept of networks and networking has been described. Further different types of networks have been discussed with practical examples. Some of the basic terminologies have been defined. At the end of each section of this unit self assessment questions have been given to evaluate the understanding.

1.5 Review Questions

Basic Network Types

Matching

Colleges and universities often have campuses at several different locations. Each campus typically has several buildings; examples include the science building, the library, the administration building, and so on. Within each of the buildings are various departments and workgroups. The computer and biology departments have several labs, the administration has a dean's office, an accounting group, and so on. As a rule, university systems are networked to the Internet. Using this information, complete the matching questions.

1. ---	Wide Area Network	A.	The biology lab
2. ---	Metropolitan Area Network	B.	The Internet
3. ---	Peer-to-Peer	C.	The south and north campuses of a college
4. ---	Client/Server	D.	The science department

True/False

1. ---	A metropolitan Area Network is usually found in one building.
2. ---	Two computers connected together are considered a network.
3. ---	The most common type of network in small organization is a WAN.
4. ---	Peer-to-peer LANs connect 1000s of computer devices.
5. ---	An enterprise network is owned and operated by the corporation using the network.

Short Essay

- Name and define the two Local Area Network types.
- What is the difference between a MAN and a WAN? Give an example of each type.

Unit II – Network Topologies

What You Will Learn (Objectives)

After completing this unit you will be able to:

- Diagram bus, star, ring, and hybrid topologies.
- Compare and contrast the advantages and disadvantages of bus, star, ring, and hybrid topologies.
- Select the appropriate topology for a particular set of network requirements.
- Create a questionnaire for an administrator to use when planning or expanding a network.

2.1 Network Topologies

Physical geographical connectivity of the network is defined as the topology of the network.

“Physical topology” is simply where the workstations and cable are physically placed. The “logical” topology is not where the devices are physically positioned, it is the actual path the *data signal* takes when transmitted. Here in this section we will talk about the physical topology. You can see where workstations are located, but you cannot see the route taken by the data. When planning a reliable network, administrators must consider several factors prior to selecting the physical and logical topology. Such factors include, ease of maintenance and management, cost, traffic, security, reliability, and redundancy. There are several topology choices, including bus, ring, star, and hybrid topology. These topologies are regular topologies. Mesh is an example of irregular topology.

2.1.1 Terminology

1. Active Hub- Type of hub that monitors, amplifies, and regenerates signals. Signals are strengthened in active hubs.
2. Attenuation- The weakening of a signal as it travels over connection media; also referred to as signal degradation.
3. Bus Topology- Network topology where computer devices are connected in a row to a continuous length of cable segment. Each end of the cable segment must be terminated by means of a terminating resistor.
4. Logical Topology- Logical topology describes the actual path of data signals through a network. It does not refer to the physical layout of the network.
5. Passive Hub- Type of hub that does not take an active role in maintaining, processing, or regenerating signals. A passive hub serves only as a physical connection point for computer devices.
6. Physical Topology- The attributes and physical setup are not logical topology, of a network. Physical topology describes the physical connections and arrangement of the internetworking devices.
7. Redundancy- Networking redundancy refers to the additional equipment or measures taken to ensure continuous operation of a network. One example is

having an extra server available in case one goes down.

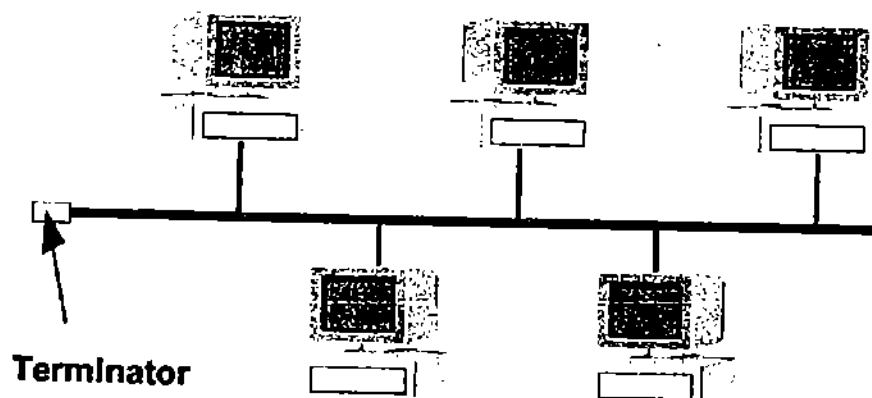
8. Ring Topology- Closed network with no beginning or end point. Computer devices are all connected to one main cable segment in a continuous fashion. Termination devices are not necessary. All computers have equal access to the network.

9. Star Topology- Topology that uses point-to-point wiring. There is a central hub, which receives and transmits signals over the network. Each computer device is connected to a hub, not directly to the other computers.

10. Terminator- Device connected at the end of each wire segment in bus networks. Terminators absorb transmission signals, which prevent them from bouncing back and causing interference.

2.2 Bus Topology

When computers were first networked together, they were simply connected to one cable segment in a series. This physical setup is called bus topology. In bus topology, data signals travel the entire length of the cable from device to device. Each end of the cable is terminated, thus preventing signal bounce back. Data signals are transmitted to the entire network and devices can send data at any time. Small networks do well with this topology, but problems increase significantly when the network becomes too large.



Bus topology advantages:

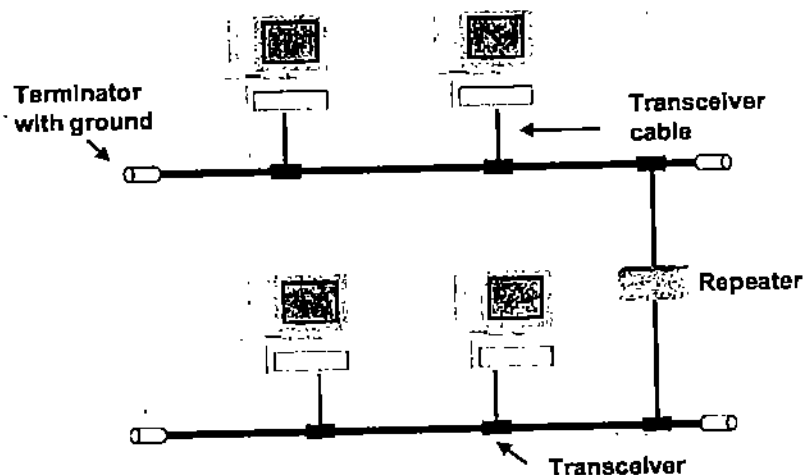
- Inexpensive to maintain
- Requires less cable
- Does not require extensive training
- Good choice for small networks

Bus topology disadvantages:

- Difficult to isolate malfunctions because of series connections
- When one device fails, all devices fail
- Heavy traffic causes considerable slowdowns and network crashes

Bus topology requires less cable than other topologies since it is a continuous series and not a point-to-point network. This keeps the costs down. Extending bus topology is accomplished by joining two cable segments with a connector or by adding repeaters. This can create transmission delays and errors. Technically speaking, bus networks are easy to use and understand and do not require extensive training. As networks grow, the excess traffic can slow the network or cause it to crash. All of these factors made it an excellent topology choice for small workplace and home networks in the preliminary stages of networking.

Bus Topology Extended With Repeater



Troubleshooting bus topology is complicated because it is difficult to isolate problems. When one device malfunctions, that device can cause an entire network to fail. Expansion also creates problems. When making modifications to the network, the entire network must be disabled. Bus topology can make maintenance and troubleshooting challenging because it is difficult to isolate problems. This, in effect, cancels the inexpensive advantage. Although the easiest and least expensive, bus topology is not practical in large, multi-room, multi-floor, multi-building installations where frequent interruptions in service might be necessary and could cause considerable problems.

Of the various topology schemes, it uses the least amount of cable, because it is a continuous and not a point-to-point topology. If you are looking for a small inexpensive, easy to understand network, bus topology may be the answer.

2.3 Ring Topology

Like bus topology, all computer devices are connected to the same cable segment. However, it is one continuous connection with no beginning or end point. Termination is not required. The signal flows in only one direction in ring topology and each device in the ring receives the signal and examines it. If the transmission is not intended for that device, the signal is regenerated and passed on to the next device in the ring.

Ring topology advantages:

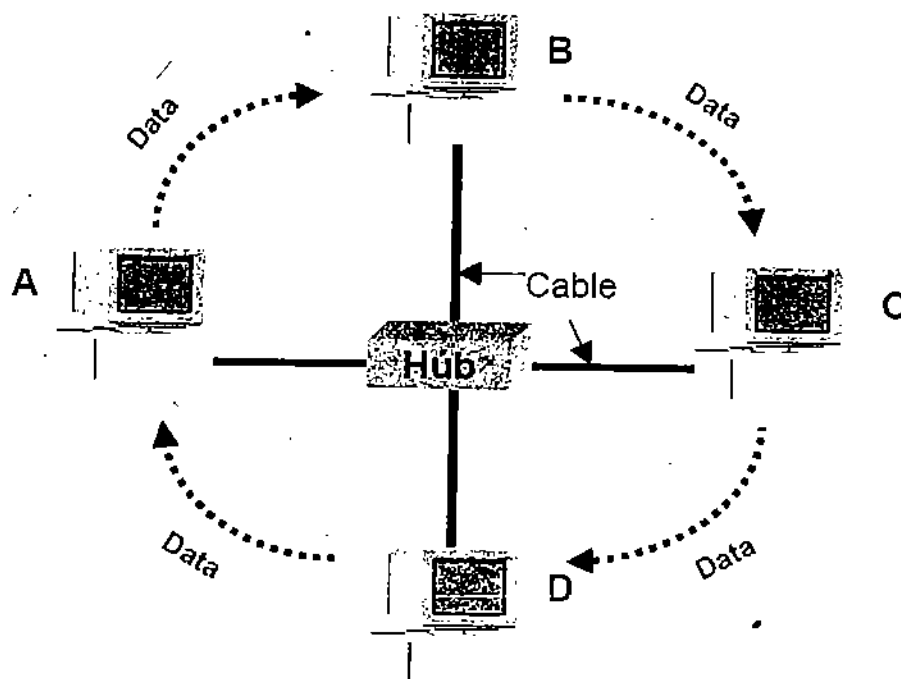
- Provides equal access for all devices on the ring

- Easier to manage and maintain than bus topology
- Very reliable
- Handles high volume traffic well

Ring topology disadvantages:

- Difficult to isolate malfunctions
- Expansion of network disrupts services for all

Each device has equal access to the network and is guaranteed access at regular intervals. This is important in a business where regular movement of data is essential, such as checks or other banking transactions.



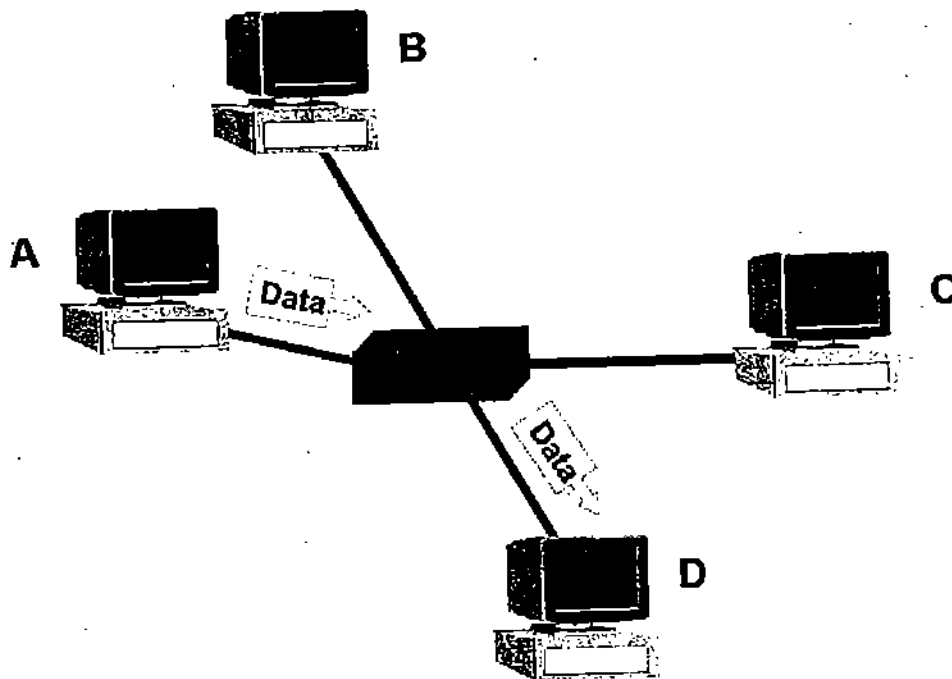
In ring topology, a token travels from computer to computer until it reaches a node waiting to transmit data. The data then attaches to the token and is delivered to the receiving device. The token then continues around the ring looking for another device waiting to send data.

Ring topology is easier to manage and maintain than bus topology, is more reliable than a bus, and it handles traffic well. On the negative side, it does require more cabling than a bus, especially if the computer devices are far apart. Expansion of a ring network disrupts data transmission for all.

If your network has high traffic, and you are looking for a reliable network where each workstation has equal access, ring topology may be the answer. Token Ring and Fiber Distributed Data Interface (FDDI) networks use token passing ring topologies.

2.4 Star Topology

Star topology is a point-to-point architectural design where all computer devices are connected to a central hub, through which all data signals must travel. There are both active and passive hubs. Passive hubs send data without amplification; active hubs amplify data signals.



Star topology advantages:

- Easy to install and upgrade
- Easier to manage and maintain than bus and ring topologies
- Central hub makes troubleshooting easy

Star topology disadvantages:

- Hub failure causes entire network to go down
- More expensive for cabling

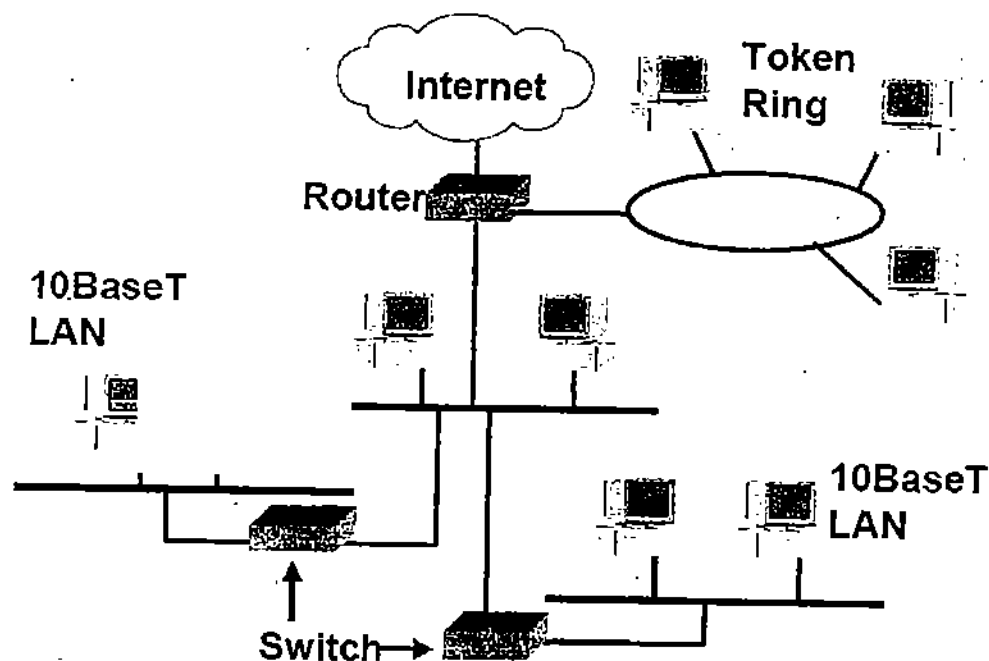
In star networks, the logical transmittal of data is similar to bus topology but only one computer can transmit data at a time. The physical topologies are quite

different. The physical hub of the star acts to logically connect all devices as if to a single cable segment. Star is the easiest topology to install, upgrade, and manage. Computer devices are attached directly to central hubs via patch cable. Each patch cable connects to a port on the hub. Small hubs usually have 4, 8, or 16 ports. Larger hubs may have up to 512 ports. Cabling is more expensive than for other topologies since each device must be connected directly to the hub

When a device on a star topology network fails, it does not disrupt the other computer devices. Similarly, if you add a new device, service to other nodes continues uninterrupted. However, if the hub fails, the entire network goes down. 10BASE-T is one of the most popular star topology networks.

2.5 Hybrid Network Topology

Hybrid topology is any combination of bus, star, and ring topology, for example, a star-bus configuration. With a star-bus network, several hubs can be connected on a bus segment to several star topology segments.



More often than not, hybrid topology is commonplace in large networks because it allows combining sub networks, each employing the least expensive, most efficient topology.

Check Your Understanding

- List several factors an administrator must consider when planning a network.
- Why must cable segments on a bus network be terminated?

- What are some advantages of bus topology?
- What is the major disadvantage of star topology?

2.6 Network Topologies

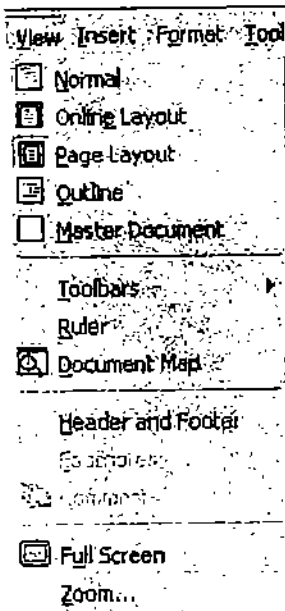
In the connectivity lesson, you diagramed all the devices and workstations connected to your network. When administrators design network topologies, they often use software applications to illustrate the topology. Using a computer, generate designs for four different network topologies, bus, ring, star, and hybrid topologies. Print out your networks and save in your portfolio.

Materials Needed

- Word processing application (e.g., MS Word)
- Network diagram from Lesson 2
- Actual network design from your instructor

Open a word processing application. Directions shown here are for Microsoft Word 6.0. (If you use another word processing application, check with your instructor for vendor specific directions.)

1. In the toolbar, click View. Scroll down to Toolbars and check Drawing.



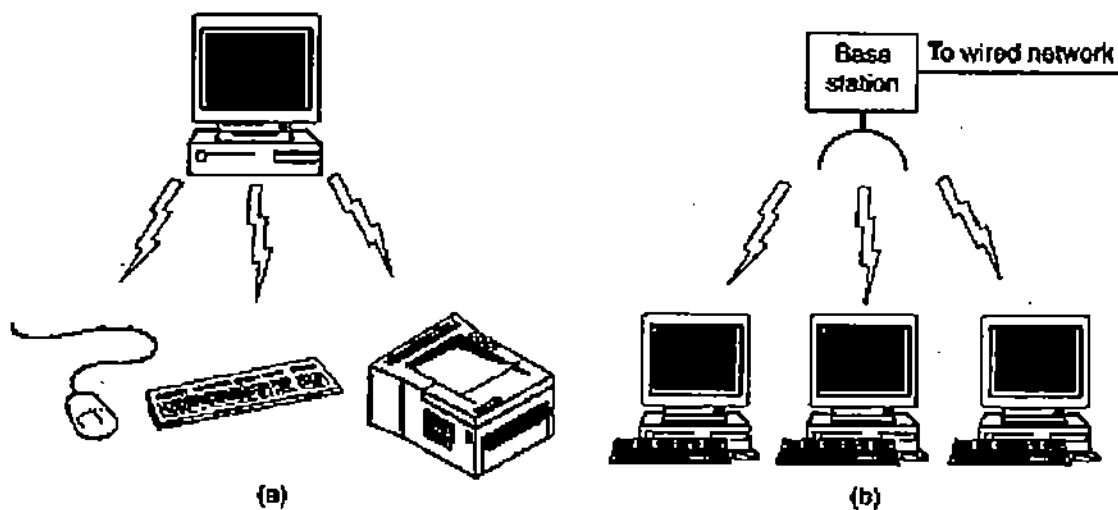
2. The following drawing toolbar will appear on the screen.



computer vendors offer the option of sending a technician to the user's home to do it. Consequently, some companies got together to design a short-range wireless network called Bluetooth to connect these components without wires. Bluetooth also allows digital cameras, headsets, scanners, and other devices to connect to a computer by merely being brought within range. No cables, no driver installation, just put them down, turn them on, and they work. For many people, this ease of operation is a big plus.

In the simplest form, system interconnection networks use the master-slave paradigm of the figure. The system unit is normally the master, talking to the mouse, keyboard, etc., as slaves. The master tells the slaves what addresses to use, when they can broadcast, how long they can transmit, what frequencies they can use, and so on.

Figure (a) Bluetooth configuration. (b) Wireless LAN.



The next step up in wireless networking are the wireless LANs. These are systems in which every computer has a radio modem and antenna with which it can communicate with other systems. Often there is an antenna on the ceiling that the machines talk to, as shown in figure. However, if the systems are close enough, they can communicate directly with one another in a peer-to-peer configuration. Wireless LANs are becoming increasingly common in small offices and homes, where installing Ethernet is considered to cause too much trouble, as well as in older office buildings, company cafeterias, conference rooms, and other places. There is a standard for wireless LANs, called IEEE 802.11, which most systems implement and which is becoming very widespread.

The third kind of wireless network is used in wide area systems. The radio network used for cellular telephones is an example of a low-bandwidth wireless system. This system has already gone through three generations. The first generation was analog and for voice only. The second generation was digital and for voice only. The third generation is digital and is for both voice and data. In a certain sense, cellular wireless networks are like wireless LANs, except that the distances involved are much greater and the bit rates much lower. Wireless LANs can operate at rates up to about 50 Mbps over distances of

2.7 Star Network

In this lesson, you learned the different types of wiring topology for networks. You have to be connected to the network in order to transmit data. In the connectivity lesson, you installed a NIC. In this activity, you will design a star and bus network topology and experience what happens when a network connection is broken.

Materials Needed

- Cables
- Networking equipment, e.g. a hub and several computers
- Computer file enabled for sharing

Star Network

Work in groups of four and set up a star network.

1. Verify that all four nodes are securely connected to the hub in a star format.
2. Check the NIC connection on each workstation to be sure there is a secure connection. Have your instructor check your star network.
3. Use one of the nodes as a server and the other three nodes as clients. Your instructor will tell you which computer will serve as the server and which will be the clients.
4. Clients should access a folder from the server. Make sure that you are able to access and use the shared folder as specified by your instructor.
5. Leave the folder open and remove the NIC connection from the server.
 - Record in the space below what happens.
 - Explain why this happened. Are the other devices able to stay connected? Why?
6. Reconnect the computer to the network.
 - Are you still connected to the shared file?

2.8 Network Wizards Network Design

Materials Needed

- Spreadsheet application

- Drawing application
- Cable cost analysis prepared for Stretch Yourself, Lesson 3

Work with a partner to create a list of questions you should ask a client who has hired you to build a network. The client wants to know which type of cable you plan to use and why, how much the cable will cost, and why you selected the particular topology. The client also wants you to submit a proposal that defends your choice of cabling and network design.

After you create your questionnaire, role-play with your partner. One of you will assume the role of client, and the other the role of the network designer. The client will answer all the questions for a fictitious company of his/her choice. The network designer will record all of the answers.

When the questionnaire is complete, you will both analyze the requirements and determine the appropriate topology. Create a computer diagram of the proposed network for the client. Prepare a spreadsheet that shows a cost analysis of the cable needed to complete the job. The cost analysis should include the type of cable, length of cable, cost of each type of cable, and the total cable cost. Use the cost information you gathered in the cabling lesson.

Submit a proposal to the client indicating the suggested network topology. Include a summary to convince the client that this is the best, most cost efficient design for her/his needs. Your proposal must also defend your cable choices. Attach the computer diagram and cable cost analysis spreadsheet to your proposal.

2.9 Summery

In this unit detail description of the popular types of topologies used in computer networks has been described. Some of the important terminologies related with the topology have been defined. Those terms have been described in the corresponding sections. At the end of each section questions for understanding have been given.

2.10 Review Questions

Part A

1. Diagram a bus network topology.
2. Diagram a star network topology.
3. Diagram a ring network topology.

Part B

1. List the advantages and disadvantages of bus network topology.
2. List the advantages and disadvantages of star topology.
3. List the advantages and disadvantages of ring topology.

Part C

1. You have been asked to create a network for five computers. Low cost is important and it is unlikely that the network will expand. Which network topology would you recommend and why?
2. Which topology would you recommend for a network that will be reconfigured frequently, must be very reliable, and is easy to troubleshoot without bringing down the entire network? Why?
3. You are designing a network for an automobile factory. There are lots of motors and fluorescent lights. Cost is not a factor. Each of the workstations must have equal access to the network. Which topology would you choose? Why?

Unit III - Network Protocols, Hardware and Software

What You Will Learn (Objectives)

After completing this unit you will be able to:

- Define Network Protocols and Standards.
- Describe Network Hardware.
- Describe Network Software.

3.1 INTRODUCTION

The exchange of information between computers for the purpose of resource sharing is generally referred to as computer communication. In order to communicate effectively, there must be a high degree of cooperation between the two computer systems. Thus, in discussing computer communications, the concept of protocol is of utmost significance. Since computer communications involve heterogeneity of hardware and software platforms, it is also important to adopt standards for implementing protocols. This unit discusses these concepts, along with some international and *de facto* standards.

3.2 Networking Protocols

When computers, terminals, and other data processing devices exchange data, the procedures involved can be quite complex. Consider, for example, the transfer of a file between two computers, either directly via point to point link or indirectly via a communication network. But more is needed. These are the typical tasks to be performed:

1. The source system must either activate the direct communication path or inform the communication network of the identity of the desired destination system.
2. The source system must ascertain that the destination system is prepared to receive data.
3. The file transfer application on the source system must ascertain that the file management program on the destination system is prepared to accept and store the file.

A protocol is used for communication between entities in different systems. The terms 'entity' and 'system' are used in very general sense. Examples of entities are user application programs, file transfer packages, database management systems, computers, terminals, and remote sensors. Note that in some cases the entity and the system in which it resides are coexistent (e.g. terminals). In general, an entity is anything capable of sending or receiving information and a system is a physically distinct object that contains one or more entities. A protocol performs the following functions:

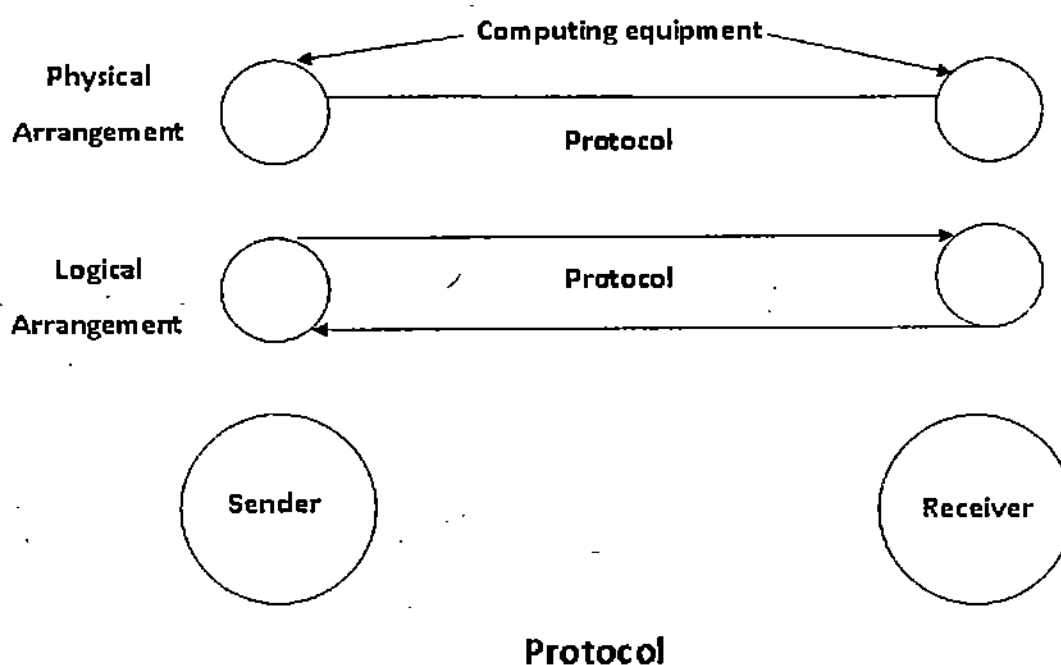
1. Establishes necessary connections.
2. Establishes a standard communication path.
3. Establishes a standard data element.

For two entities to communicate successfully, they must speak the same language. What is communicated, how it is communicated, and when it is communicated must conform to some conventions mutually acceptable to the entities involved. The conventions are referred to as

'protocol', which may be defined as a set of rules governing the exchange of data between two entities. The key elements of a protocol are:

- **Syntax:** includes things such as data format and signal levels.
- **Semantics:** includes control information for coordination and error handling
- **Timing:** includes speed matching and sequencing

Having introduced the concept of a protocol, we can now introduce the concept of computer communications architecture. We make the observation that the task of communicating between two entities on different systems is too complicated to be handled by a single process or module. For example, when two file transfer packages exchange data, each invokes a network services module which performs the task of exchanging data with its local file transfer module. The local file transfer module employs a system-to-system protocol. To actually transfer the data, the network services module invokes the station-node services module which exchanges data with the next immediate node on the communication network. Thus, instead of a single protocol, there is a structured set of protocols that implement the communication function. This structure is referred to as computer-communication architecture.



3.3 Standards

Standards are required to govern the physical, electrical, and procedural characteristics of communication equipment.

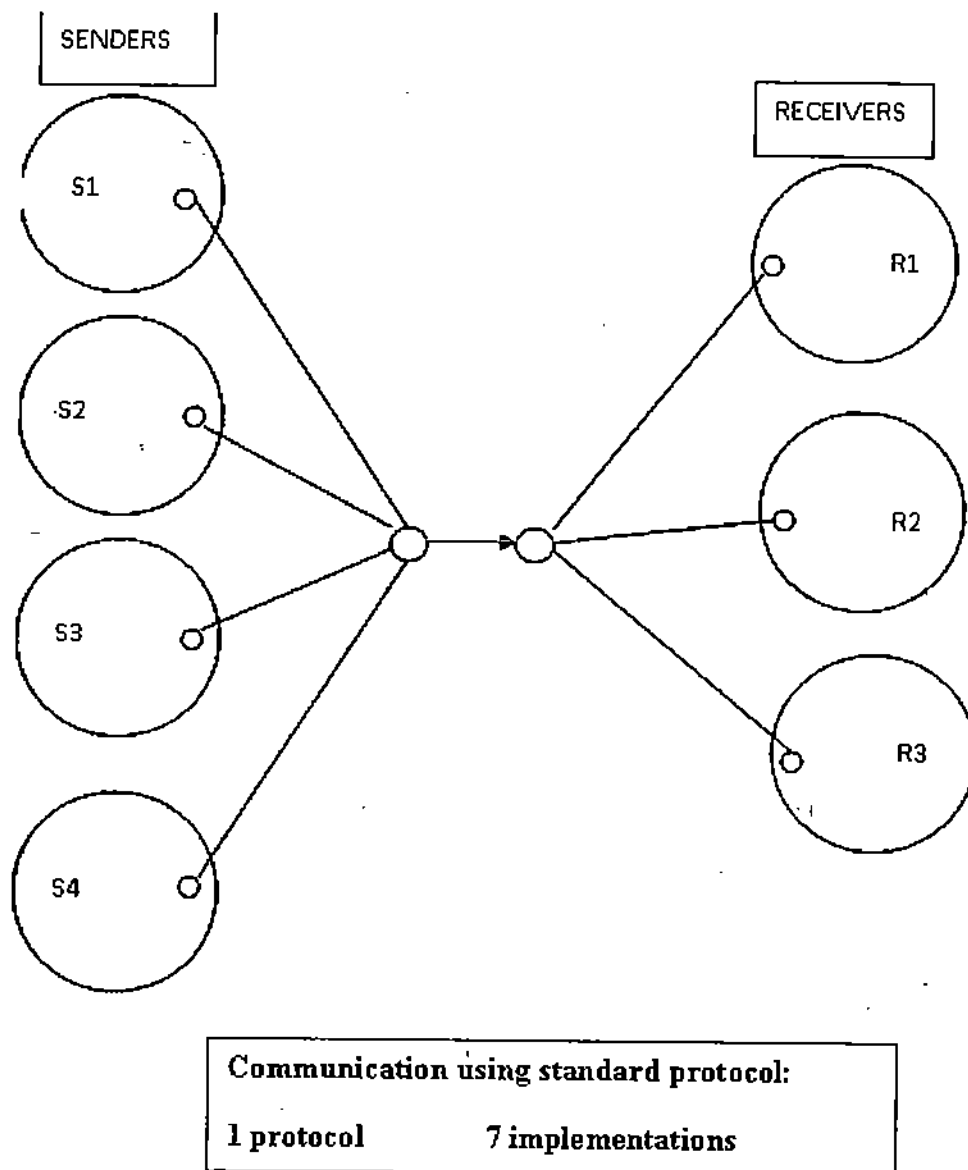
'A standard is a prescribed set of rules, conditions, or requirements concerning definition of terms; classification of components; specification of materials, performance, or operations; delineation of procedures; or measurement of quantity and quality in describing materials, products, systems, services or practices'.

The principal disadvantage of standards is that they tend to freeze technology. By the time a standard is developed, subjected to review and communication with others, computers vendors would have attempted to monopolize their customers. The proliferation of computers and distributed processing has made that an unattainable position. Computers from different vendors must communicate with each other. With the ongoing evolution of protocol standards, customers will no longer accept special purpose protocol conversion software development. The day is quickly coming when the standards discussed in this text will dominate the market.

The key advantages of standardization are:

1. A standard assures that there will be a large market for a particular piece of equipment or software. This encourages mass production, the use of LSI and VLSI techniques, resulting in lower costs.
2. A standard allows products from multiple vendors to communicate with each other, giving the purchaser more flexibility in equipment selection and use.

Figure shows a scenario in which communication is achieved with a standard.



3.4 Network Hardware

It is now time to turn our attention from the applications and social aspects of networking (the fun stuff) to the technical issues involved in network design (the work stuff). There is no generally accepted taxonomy into which all computer networks fit, but two dimensions stand out as important: transmission technology and scale. We will now examine each of these in turn.

Broadly speaking, there are two types of transmission technology that are in widespread use. They are as follows:

1. Broadcast links.
2. Point-to-point links.

Broadcast networks have a single communication channel that is shared by all the machines on the network. Short messages, called packets in certain contexts, sent by any machine are received by all the others. An address field within the packet specifies the intended recipient. Upon receiving a packet, a machine checks the address field. If the packet is intended for the receiving machine, that machine processes the packet; if the packet is intended for some other machine, it is just ignored.

As an analogy, consider someone standing at the end of a corridor with many rooms off it and shouting "Watson, come here. I want you." Although the packet may actually be received (heard) by many people, only Watson responds. The others just ignore it. Another analogy is an airport announcement asking all flight 644 passengers to report to gate 12 for immediate boarding.

Broadcast systems generally also allow the possibility of addressing a packet to all destinations by using a special code in the address field. When a packet with this code is transmitted, it is received and processed by every machine on the network. This mode of operation is called broadcasting. Some broadcast systems also support transmission to a subset of the machines, something known as multicasting. One possible scheme is to reserve one bit to indicate multicasting. The remaining $n - 1$ address bits can hold a group number. Each machine can "subscribe" to any or all of the groups. When a packet is sent to a certain group, it is delivered to all machines subscribing to that group.

In contrast, point-to-point networks consist of many connections between individual pairs of machines. To go from the source to the destination, a packet on this type of network may have to first visit one or more intermediate machines. Often multiple routes, of different lengths, are possible, so finding good ones is important in point-to-point networks. As a general rule (although there are many exceptions), smaller, geographically localized networks tend to use broadcasting, whereas larger networks usually are point-to-point. Point-to-point transmission with one sender and one receiver is sometimes called unicasting.

An alternative criterion for classifying networks is their scale. In following figure we classify multiple processor systems by their physical size. At the top are the personal area networks, networks that are meant for one person. For example, a wireless network connecting a computer

networks that are meant for one person. For example, a wireless network connecting a computer with its mouse, keyboard, and printer is a personal area network. Also, a PDA that controls the user's hearing aid or pacemaker fits in this category. Beyond the personal area networks come longer-range networks. These can be divided into local, metropolitan, and wide area networks. Finally, the connection of two or more networks is called an internetwork. The worldwide Internet is a well-known example of an internetwork. Distance is important as a classification metric because different techniques are used at different scales. In this book we will be concerned with networks at all these scales. Below we give a brief introduction to network hardware.

Figure: Classification of interconnected processors by scale

Interprocessor distance	Processors located in same	Example
1 m	Square meter	Personal area network
10 m	Room	
100 m	Building	Local area network
1 km	Campus	
10 km	City	Metropolitan area network
100 km	Country	
1000 km	Continent	Wide area network
10,000 km	Planet	
		The Internet

3.4.1 Wireless Networks

Digital wireless communication is not a new idea. As early as 1901, the Italian physicist Guglielmo Marconi demonstrated a ship-to-shore wireless telegraph, using Morse Code (dots and dashes are binary, after all). Modern digital wireless systems have better performance, but the basic idea is the same.

To a first approximation, wireless networks can be divided into three main categories:

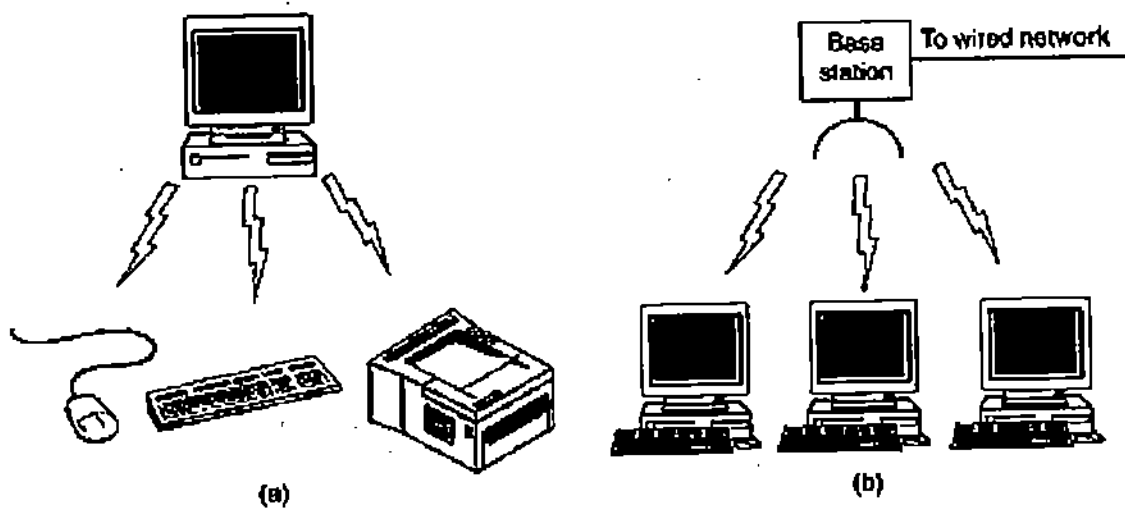
1. System interconnection.
2. Wireless LANs.
3. Wireless WANs.

System interconnection is all about interconnecting the components of a computer using short-range radio. Almost every computer has a monitor, keyboard, mouse, and printer connected to the main unit by cables. So many new users have a hard time plugging all the cables into the right little holes (even though they are usually color coded) that most

computer vendors offer the option of sending a technician to the user's home to do it. Consequently, some companies got together to design a short-range wireless network called Bluetooth to connect these components without wires. Bluetooth also allows digital cameras, headsets, scanners, and other devices to connect to a computer by merely being brought within range. No cables, no driver installation, just put them down, turn them on, and they work. For many people, this ease of operation is a big plus.

In the simplest form, system interconnection networks use the master-slave paradigm of the figure. The system unit is normally the master, talking to the mouse, keyboard, etc., as slaves. The master tells the slaves what addresses to use, when they can broadcast, how long they can transmit, what frequencies they can use, and so on.

Figure (a) *Bluetooth configuration.* (b) *Wireless LAN.*



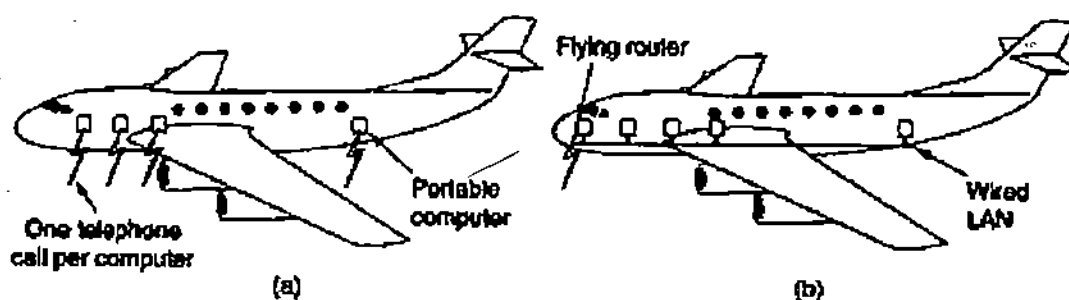
The next step up in wireless networking are the wireless LANs. These are systems in which every computer has a radio modem and antenna with which it can communicate with other systems. Often there is an antenna on the ceiling that the machines talk to, as shown in figure. However, if the systems are close enough, they can communicate directly with one another in a peer-to-peer configuration. Wireless LANs are becoming increasingly common in small offices and homes, where installing Ethernet is considered to cause too much trouble, as well as in older office buildings, company cafeterias, conference rooms, and other places. There is a standard for wireless LANs, called IEEE 802.11, which most systems implement and which is becoming very widespread.

The third kind of wireless network is used in wide area systems. The radio network used for cellular telephones is an example of a low-bandwidth wireless system. This system has already gone through three generations. The first generation was analog and for voice only. The second generation was digital and for voice only. The third generation is digital and is for both voice and data. In a certain sense, cellular wireless networks are like wireless LANs, except that the distances involved are much greater and the bit rates much lower. Wireless LANs can operate at rates up to about 50 Mbps over distances of

tens of meters. Cellular systems operate below 1 Mbps, but the distance between the base station and the computer or telephone is measured in kilometers rather than in meters.

In addition to these low-speed networks, high-bandwidth wide area wireless networks are also being developed. The initial focus is high-speed wireless Internet access from homes and businesses, bypassing the telephone system. This service is often called local multipoint distribution service. We will study it later in the book. A standard for it, called IEEE 802.16, has also been developed.

Almost all wireless networks hook up to the wired network at some point to provide access to files, databases, and the Internet. There are many ways these connections can be realized, depending on the circumstances. For example, in Figure (a), we depict an airplane with a number of people using modems and seat-back telephones to call the office. Each call is independent of the other ones. A much more efficient option, however, is the flying LAN of Figure (b). Here each seat comes equipped with an Ethernet connector into which passengers can plug their computers. A single router on the aircraft maintains a radio link with some router on the ground, changing routers as it flies along. This configuration is just a traditional LAN, except that its connection to the outside world happens to be a radio link instead of a hardwired line.



3.4.2 Home Networks

Home networking is on the horizon. The fundamental idea is that in the future most homes will be set up for networking. Every device in the home will be capable of communicating with every other device, and all of them will be accessible over the Internet. This is one of those visionary concepts that nobody asked for (like TV remote controls or mobile phones), but once they arrived nobody can imagine how they lived without them.

Many devices are capable of being networked. Some of the more obvious categories (with examples) are as follows:

1. Computers (desktop PC, notebook PC, PDA, shared peripherals).
2. Entertainment (TV, DVD, VCR, camcorder, camera, stereo, MP3).
3. Telecommunications (telephone, mobile telephone, intercom, fax).
4. Appliances (microwave, refrigerator, clock, furnace, airco, lights).

5. Telemetry (utility meter, smoke/burglar alarm, thermostat, babycam).

Home computer networking is already here in a limited way. Many homes already have a device to connect multiple computers to a fast Internet connection. Networked entertainment is not quite here, but as more and more music and movies can be downloaded from the Internet, there will be a demand to connect stereos and televisions to it. Also, people will want to share their own videos with friends and family, so the connection will need to go both ways. Telecommunications gear is already connected to the outside world, but soon it will be digital and go over the Internet. The average home probably has a dozen clocks (e.g., in appliances), all of which have to be reset twice a year when daylight saving time (summer time) comes and goes. If all the clocks were on the Internet, that resetting could be done automatically. Finally, remote monitoring of the home and its contents is a likely winner. Probably many parents would be willing to spend some money to monitor their sleeping babies on their PDAs when they are eating out, even with a rented teenager in the house. While one can imagine a separate network for each application area, integrating all of them into a single network is probably a better idea.

Home networking has some fundamentally different properties than other network types. First, the network and devices have to be easy to install. The author has installed numerous pieces of hardware and software on various computers over the years, with mixed results. A series of phone calls to the vendor's helpdesk typically resulted in answers like (1) Read the manual, (2) Reboot the computer, (3) Remove all hardware and software except ours and try again, (4) Download the newest driver from our Web site, and if all else fails, (5) Reformat the hard disk and then reinstall Windows from the CD-ROM. Telling the purchaser of an Internet refrigerator to download and install a new version of the refrigerator's operating system is not going to lead to happy customers. Computer users are accustomed to putting up with products that do not work; the car-, television-, and refrigerator-buying public is far less tolerant. They expect products to work for 100% from the word go.

Second, the network and devices have to be foolproof in operation. Air conditioners used to have one knob with four settings: OFF, LOW, MEDIUM, and HIGH. Now they have 30-page manuals. Once they are networked, expect the chapter on security alone to be 30 pages. This will be beyond the comprehension of virtually all the users.

Third, low price is essential for success. People will not pay a \$50 premium for an Internet thermostat because few people regard monitoring their home temperature from work that important. For \$5 extra, it might sell, though.

Fourth, the main application is likely to involve multimedia, so the network needs sufficient **capacity**. There is no market for Internet-connected televisions that show shaky movies at 320 x 240 pixel resolution and 10 frames/sec. Fast Ethernet, the workhorse in most offices, is not good enough for multimedia. Consequently, home networks will need better performance than that of existing office networks and at lower prices before they become mass market items.

Fifth, it must be possible to start out with one or two devices and expand the reach of the network gradually. This means no format wars. Telling consumers to buy peripherals with IEEE 1394 (FireWire) interfaces and a few years later retracting that and saying USB 2.0 is the interface-of-the-month is going to make consumers skittish. The network interface will have to remain stable for many years; the wiring (if any) will have to remain stable for decades.

Sixth, security and reliability will be very important. Losing a few files to an e-mail virus is one thing; having a burglar disarm your security system from his PDA and then plunder your house is something quite different.

An interesting question is whether home networks will be wired or wireless. Most homes already have six networks installed: electricity, telephone, cable television, water, gas, and sewer. Adding a seventh one during construction is not difficult, but retrofitting existing houses is expensive. Cost favors wireless networking, but security favors wired networking. The problem with wireless is that the radio waves they use are quite good at going through fences. Not everyone is overjoyed at the thought of having the neighbors piggybacking on their Internet connection and reading their e-mail on its way to the printer.

In short, home networking offers many opportunities and challenges. Most of them relate to the need to be easy to manage, dependable, and secure, especially in the hands of nontechnical users, while at the same time delivering high performance at low cost.

3.4.3 Internetworks

Many networks exist in the world, often with different hardware and software. People connected to one network often want to communicate with people attached to a different one. The fulfillment of this desire requires that different, and frequently incompatible networks, be connected, sometimes by means of machines called gateways to make the connection and provide the necessary translation, both in terms of hardware and software. A collection of interconnected networks is called an internetwork or internet. These terms will be used in a generic sense, in contrast to the worldwide Internet (which is one specific internet), which we will always capitalize.

A common form of internet is a collection of LANs connected by a WAN. The only real technical distinction between a subnet and a WAN in this case is whether hosts are present. If the system within the gray area contains only routers, it is a subnet; if it contains both routers and hosts, it is a WAN. The real differences relate to ownership and use.

Subnets, networks, and internetworks are often confused. Subnet makes the most sense in the context of a wide area network, where it refers to the collection of routers and communication lines owned by the network operator. As an analogy, the telephone system consists of telephone switching offices connected to one another by high-speed lines, and to houses and businesses by low-speed lines. These lines and equipment, owned and managed by the telephone company, form the subnet of the telephone system. The telephones themselves (the hosts in this analogy) are not part of the subnet. The

combination of a subnet and its hosts forms a network. In the case of a LAN, the cable and the hosts form the network. There really is no subnet.

An internetwork is formed when distinct networks are interconnected. In our view, connecting a LAN and a WAN or connecting two LANs forms an internetwork, but there is little agreement in the industry over terminology in this area. One rule of thumb is that if different organizations pay to construct different parts of the network and each maintains its part, we have an internetwork rather than a single network. Also, if the underlying technology is different in different parts (e.g., broadcast versus point-to-point), we probably have two networks.

3.5 Network Software

The first computer networks were designed with the hardware as the main concern and the software as an afterthought. This strategy no longer works. Network software is now highly structured. In the following sections we examine the software structuring technique in some detail.

3.5.1 Protocol Hierarchies

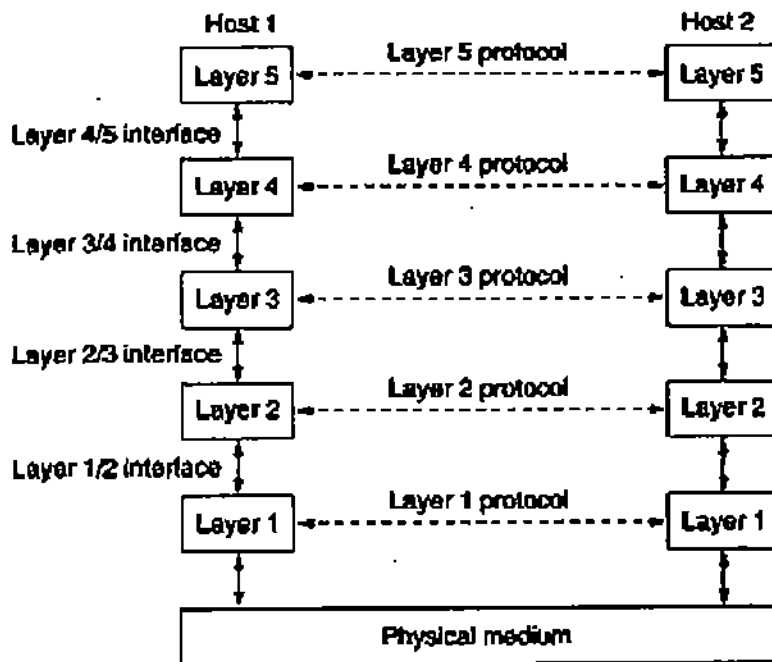
To reduce their design complexity, most networks are organized as a stack of layers or levels, each one built upon the one below it. The number of layers, the name of each layer, the contents of each layer, and the function of each layer differ from network to network. The purpose of each layer is to offer certain services to the higher layers, shielding those layers from the details of how the offered services are actually implemented. In a sense, each layer is a kind of virtual machine, offering certain services to the layer above it.

This concept is actually a familiar one and used throughout computer science, where it is variously known as information hiding, abstract data types, data encapsulation, and object-oriented programming. The fundamental idea is that a particular piece of software (or hardware) provides a service to its users but keeps the details of its internal state and algorithms hidden from them.

Layer n on one machine carries on a conversation with layer n on another machine. The rules and conventions used in this conversation are collectively known as the layer n protocol. Basically, a protocol is an agreement between the communicating parties on how communication is to proceed. As an analogy, when a woman is introduced to a man, she may choose to stick out her hand. He, in turn, may decide either to shake it or kiss it, depending, for example, on whether she is an American lawyer at a business meeting or a European princess at a formal ball. Violating the protocol will make communication more difficult, if not completely impossible.

A five-layer network is illustrated in Figure. The entities comprising the corresponding layers on different machines are called peers. The peers may be processes, hardware devices, or even human beings. In other words, it is the peers that communicate by using the protocol.

Figure: Layers, protocols, and interfaces



In reality, no data are directly transferred from layer n on one machine to layer n on another machine. Instead, each layer passes data and control information to the layer immediately below it, until the lowest layer is reached. Below layer 1 is the physical medium through which actual communication occurs. In the figure, virtual communication is shown by dotted lines and physical communication by solid lines.

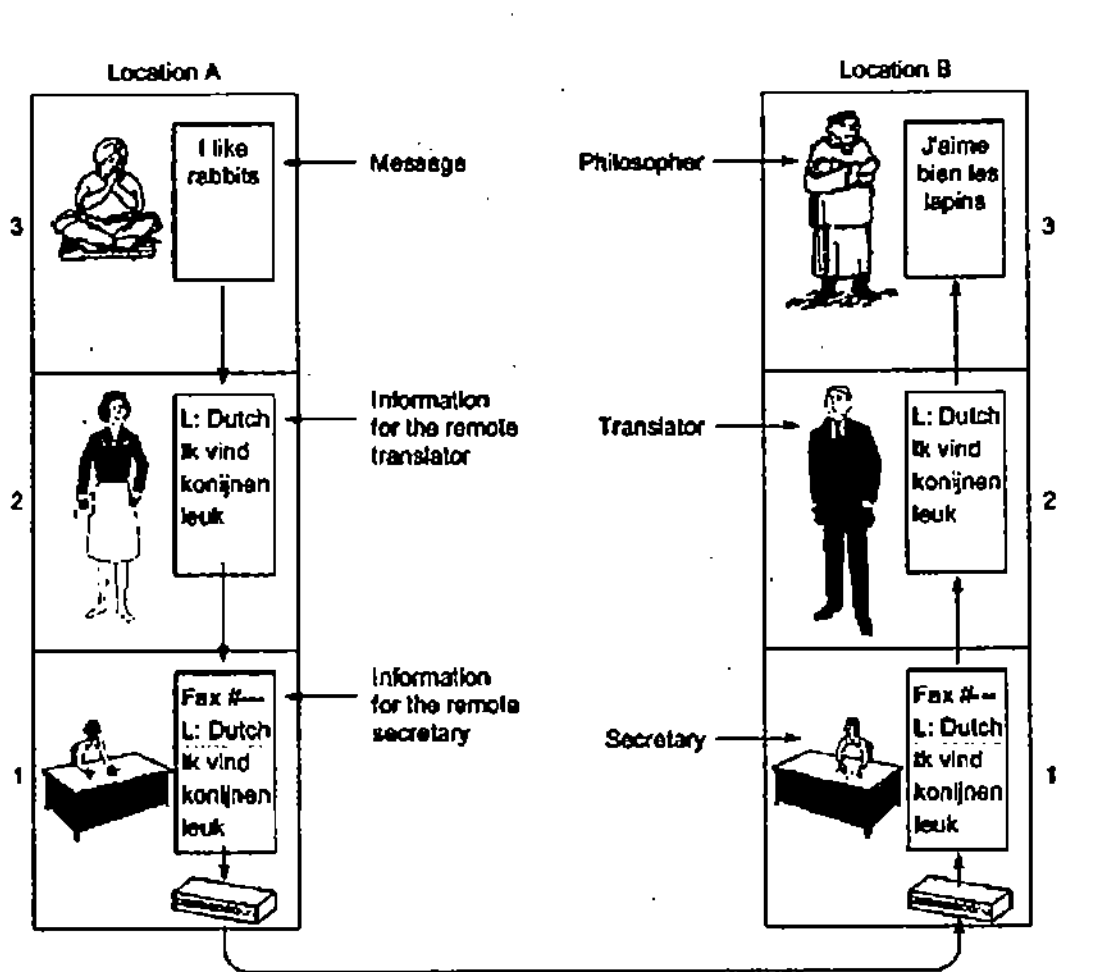
Between each pair of adjacent layers is an interface. The interface defines which primitive operations and services the lower layer makes available to the upper one. When network designers decide how many layers to include in a network and what each one should do, one of the most important considerations is defining clean interfaces between the layers. Doing so, in turn, requires that each layer perform a specific collection of well-understood functions. In addition to minimizing the amount of information that must be passed between layers, clear-cut interfaces also make it simpler to replace the implementation of one layer with a completely different implementation (e.g., all the telephone lines are replaced by satellite channels) because all that is required of the new implementation is that it offer exactly the same set of services to its upstairs neighbor as the old implementation did. In fact, it is common that different hosts use different implementations.

A set of layers and protocols is called a network architecture. The specification of an architecture must contain enough information to allow an implementer to write the program or build the hardware for each layer so that it will correctly obey the appropriate

protocol. Neither the details of the implementation nor the specification of the interfaces is part of the architecture because these are hidden away inside the machines and not visible from the outside. It is not even necessary that the interfaces on all machines in a network be the same, provided that each machine can correctly use all the protocols. A list of protocols used by a certain system, one protocol per layer, is called a protocol stack. The subjects of network architectures, protocol stacks, and the protocols themselves are the principal topics of this book.

An analogy may help explain the idea of multilayer communication. Imagine two philosophers (peer processes in layer 3), one of whom speaks Urdu and English and one of whom speaks Chinese and French. Since they have no common language, they each engage a translator (peer processes at layer 2), each of whom in turn contacts a secretary (peer processes in layer 1). Philosopher 1 wishes to convey his affection for *oryctolagus cuniculus* to his peer. To do so, he passes a message (in English) across the 2/3 interface to his translator, saying "I like rabbits," as illustrated in figure. The translators have agreed on a neutral language known to both of them, Dutch, so the message is converted to "Ik vind konijnen leuk." The choice of language is the layer 2 protocol and is up to the layer 2 peer processes.

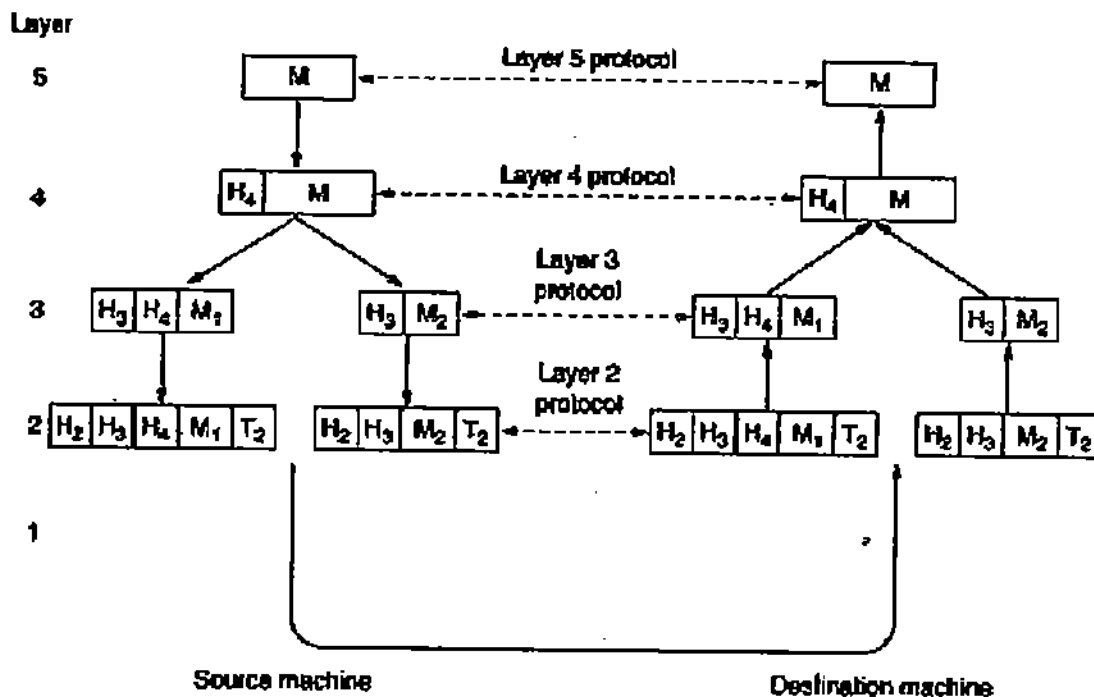
Figure: The philosopher-translator-secretary architecture



The translator then gives the message to a secretary for transmission, by, for example, fax (the layer 1 protocol). When the message arrives, it is translated into French and passed across the 2/3 interface to philosopher 2. Note that each protocol is completely independent of the other ones as long as the interfaces are not changed. The translators can switch from Dutch to say, Finnish, at will, provided that they both agree, and neither changes his interface with either layer 1 or layer 3. Similarly, the secretaries can switch from fax to e-mail or telephone without disturbing (or even informing) the other layers. Each process may add some information intended only for its peer. This information is not passed upward to the layer above.

Now consider a more technical example: how to provide communication to the top layer of the five-layer network in figure. A message, M , is produced by an application process running in layer 5 and given to layer 4 for transmission. Layer 4 puts a header in front of the message to identify the message and passes the result to layer 3. The header includes control information, such as sequence numbers, to allow layer 4 on the destination machine to deliver messages in the right order if the lower layers do not maintain sequence. In some layers, headers can also contain sizes, times, and other control fields.

Figure: Example information flow supporting virtual communication in layer 5



In many networks, there is no limit to the size of messages transmitted in the layer 4 protocol, but there is nearly always a limit imposed by the layer 3 protocol. Consequently, layer 3 must break up the incoming messages into smaller units, packets, prepending a layer 3 header to each packet. In this example, M is split into two parts, M_1 and M_2 .

Layer 3 decides which of the outgoing lines to use and passes the packets to layer 2. Layer 2 adds not only a header to each piece, but also a trailer, and gives the resulting unit to layer 1 for physical transmission. At the receiving machine the message moves upward, from layer to layer, with headers being stripped off as it progresses. None of the headers for layers below n are passed up to layer n .

The important thing to understand about above figure is the relation between the virtual and actual communication and the difference between protocols and interfaces. The peer processes in layer 4, for example, conceptually think of their communication as being "horizontal," using the layer 4 protocol. Each one is likely to have a procedure called something like Send-To-Other Side and Get-From-Other Side, even though these procedures actually communicate with lower layers across the 3/4 interface, not with the other side.

The peer process abstraction is crucial to all network design. Using it, the unmanageable task of designing the complete network can be broken into several smaller, manageable design problems, namely, the design of the individual layers.

3.5.2 Design Issues for Layers

Some of the key design issues that occur in computer networks are present in several layers. Below, we will briefly mention some of the more important ones.

Every layer needs a mechanism for identifying senders and receivers. Since a network normally has many computers, some of which have multiple processes, a means is needed for a process on one machine to specify with whom it wants to talk. As a consequence of having multiple destinations, some form of addressing is needed in order to specify a specific destination.

Another set of design decisions concerns the rules for data transfer. In some systems, data only travel in one direction; in others, data can go both ways. The protocol must also determine how many logical channels the connection corresponds to and what their priorities are. Many networks provide at least two logical channels per connection, one for normal data and one for urgent data.

Error control is an important issue because physical communication circuits are not perfect. Many error-detecting and error-correcting codes are known, but both ends of the connection must agree on which one is being used. In addition, the receiver must have some way of telling the sender which messages have been correctly received and which have not.

Not all communication channels preserve the order of messages sent on them. To deal with a possible loss of sequencing, the protocol must make explicit provision for the receiver to allow the pieces to be reassembled properly. An obvious solution is to number the pieces, but this solution still leaves open the question of what should be done with pieces that arrive out of order.

An issue that occurs at every level is how to keep a fast sender from swamping a slow receiver with data. Various solutions have been proposed and will be discussed later. Some of them involve some kind of feedback from the receiver to the sender, either directly or indirectly, about the receiver's current situation. Others limit the sender to an agreed-on transmission rate. This subject is called flow control.

Another problem that must be solved at several levels is the inability of all processes to accept arbitrarily long messages. This property leads to mechanisms for disassembling, transmitting, and then reassembling messages. A related issue is the problem of what to do when processes insist on transmitting data in units that are so small that sending each one separately is inefficient. Here the solution is to gather several small messages heading toward a common destination into a single large message and dismember the large message at the other side.

When it is inconvenient or expensive to set up a separate connection for each pair of communicating processes, the underlying layer may decide to use the same connection for multiple, unrelated conversations. As long as this multiplexing and demultiplexing is done transparently, it can be used by any layer. Multiplexing is needed in the physical layer, for example, where all the traffic for all connections has to be sent over at most a few physical circuits.

When there are multiple paths between source and destination, a route must be chosen. Sometimes this decision must be split over two or more layers. For example, to send data from London to Rome, a high-level decision might have to be made to transit France or Germany based on their respective privacy laws. Then a low-level decision might have to be made to select one of the available circuits based on the current traffic load. This topic is called routing.

3.5.3 Connection-oriented and Connectionless Services

Layers can offer two different types of service to the layers above them: connection-oriented and connectionless. In this section we will look at these two types and examine the differences between them.

Connection-oriented service is modeled after the telephone system. To talk to someone, you pick up the phone, dial the number, talk, and then hang up. Similarly, to use a connection-oriented network service, the service user first establishes a connection, uses the connection, and then releases the connection. The essential aspect of a connection is that it acts like a tube: the sender pushes objects (bits) in at one end, and the receiver takes them out at the other end. In most cases the order is preserved so that the bits arrive in the order they were sent.

In some cases when a connection is established, the sender, receiver, and subnet conduct a negotiation about parameters to be used, such as maximum message size, quality of service required, and other issues. Typically, one side makes a proposal and the other side can accept it, reject it, or make a counterproposal.

In contrast, connectionless service is modeled after the postal system. Each message (letter) carries the full destination address, and each one is routed through the system independent of all the others. Normally, when two messages are sent to the same destination, the first one sent will be the first one to arrive. However, it is possible that the first one sent can be delayed so that the second one arrives first.

Each service can be characterized by a quality of service. Some services are reliable in the sense that they never lose data. Usually, a reliable service is implemented by having the receiver acknowledge the receipt of each message so the sender is sure that it arrived. The acknowledgement process introduces overhead and delays, which are often worth it but are sometimes undesirable.

A typical situation in which a reliable connection-oriented service is appropriate is file transfer. The owner of the file wants to be sure that all the bits arrive correctly and in the same order they were sent. Very few file transfer customers would prefer a service that occasionally scrambles or loses a few bits, even if it is much faster.

Reliable connection-oriented service has two minor variations: message sequences and byte streams. In the former variant, the message boundaries are preserved. When two 1024-byte messages are sent, they arrive as two distinct 1024-byte messages, never as one 2048-byte message. In the latter, the connection is simply a stream of bytes, with no message boundaries. When 2048 bytes arrive at the receiver, there is no way to tell if they were sent as one 2048-byte message, two 1024-byte messages, or 2048 1-byte messages. If the pages of a book are sent over a network to a phototypesetter as separate messages, it might be important to preserve the message boundaries. On the other hand, when a user logs into a remote server, a byte stream from the user's computer to the server is all that is needed. Message boundaries are not relevant.

As mentioned above, for some applications, the transit delays introduced by acknowledgements are unacceptable. One such application is digitized voice traffic. It is preferable for telephone users to hear a bit of noise on the line from time to time than to experience a delay waiting for acknowledgements. Similarly, when transmitting a video conference, having a few pixels wrong is no problem, but having the image jerk along as the flow stops to correct errors is irritating.

Not all applications require connections. For example, as electronic mail becomes more common, electronic junk is becoming more common too. The electronic junk-mail sender probably does not want to go to the trouble of setting up and later tearing down a connection just to send one item. Nor is 100 percent reliable delivery essential, especially if it costs more. All that is needed is a way to send a single message that has a high probability of arrival, but no guarantee. Unreliable (meaning not acknowledged) connectionless service is often called datagram service, in analogy with telegram service, which also does not return an acknowledgement to the sender.

In other situations, the convenience of not having to establish a connection to send one short message is desired, but reliability is essential. The acknowledged datagram service can be provided for these applications. It is like sending a registered letter and requesting

a return receipt. When the receipt comes back, the sender is absolutely sure that the letter was delivered to the intended party and not lost along the way.

Still another service is the request-reply service. In this service the sender transmits a single datagram containing a request; the reply contains the answer. For example, a query to the local library asking where Uighur is spoken falls into this category. Request-reply is commonly used to implement communication in the client-server model: the client issues a request and the server responds to it. Figure summarizes the types of services discussed above.

Figure: Six different types of service

	Service		Example	
Connection-oriented	Reliable message stream		Sequence of pages	
	Reliable byte stream		Remote login	
	Unreliable connection		Digitized voice	
Connection-less	Unreliable datagram		Electronic junk mail	
	Acknowledged datagram		Registered mail	
	Request-reply		Database query	

The concept of using unreliable communication may be confusing at first. After all, why would anyone actually prefer unreliable communication to reliable communication? First of all, reliable communication (in our sense, that is, acknowledged) may not be available. For example, Ethernet does not provide reliable communication. Packets can occasionally be damaged in transit. It is up to higher protocol levels to deal with this problem. Second, the delays inherent in providing a reliable service may be unacceptable, especially in real-time applications such as multimedia. For these reasons, both reliable and unreliable communication coexist.

3.5.4 Service Primitives

A service is formally specified by a set of primitives (operations) available to a user process to access the service. These primitives tell the service to perform some action or report on an action taken by a peer entity. If the protocol stack is located in the operating system, as it often is, the primitives are normally system calls. These calls cause a trap to kernel mode, which then turns control of the machine over to the operating system to send the necessary packets.

The set of primitives available depends on the nature of the service being provided. The primitives for connection-oriented service are different from those of connectionless service. As a minimal example of the service primitives that might be provided to

implement a reliable byte stream in a client-server environment, consider the primitives listed in figure.

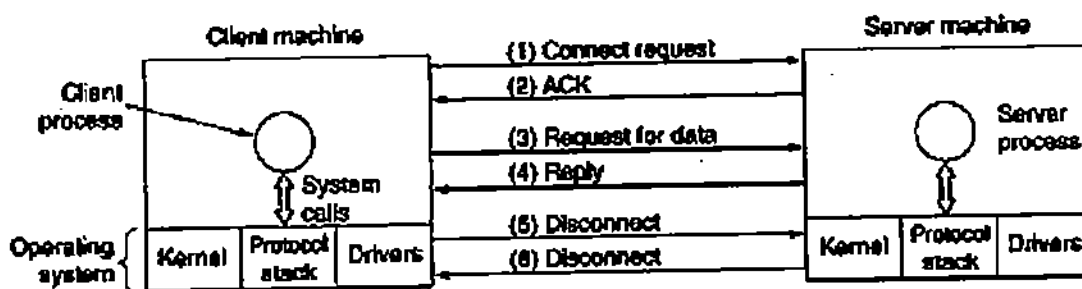
Figure: Five service primitives for implementing a simple connection-oriented service

Primitive	Meaning
LISTEN	Block waiting for an incoming connection
CONNECT	Establish a connection with a waiting peer
RECEIVE	Block waiting for an incoming message
SEND	Send a message to the peer
DISCONNECT	Terminate a connection

These primitives might be used as follows. First, the server executes LISTEN to indicate that it is prepared to accept incoming connections. A common way to implement LISTEN is to make it a blocking system call. After executing the primitive, the server process is blocked until a request for connection appears.

Next, the client process executes CONNECT to establish a connection with the server. The CONNECT call needs to specify who to connect to, so it might have a parameter giving the server's address. The operating system then typically sends a packet to the peer asking it to connect, as shown by (1) in Figure. The client process is suspended until there is a response. When the packet arrives at the server, it is processed by the operating system there. When the system sees that the packet is requesting a connection, it checks to see if there is a listener. If so, it does two things: unblocks the listener and sends back an acknowledgement (2). The arrival of this acknowledgement then releases the client. At this point the client and server are both running and they have a connection established. It is important to note that the acknowledgement (2) is generated by the protocol code itself, not in response to a user-level primitive. If a connection request arrives and there is no listener, the result is undefined. In some systems the packet may be queued for a short time in anticipation of a LISTEN.

Figure: Packets sent in a simple client-server interaction on a connection-oriented network



The obvious analogy between this protocol and real life is a customer (client) calling a company's customer service manager. The service manager starts out by being near the

telephone in case it rings. Then the client places the call. When the manager picks up the phone, the connection is established.

The next step is for the server to execute `RECEIVE` to prepare to accept the first request. Normally, the server does this immediately upon being released from the `LISTEN`, before the acknowledgement can get back to the client. The `RECEIVE` call blocks the server.

Then the client executes `SEND` to transmit its request (3) followed by the execution of `RECEIVE` to get the reply.

The arrival of the request packet at the server machine unblocks the server process so it can process the request. After it has done the work, it uses `SEND` to return the answer to the client (4). The arrival of this packet unblocks the client, which can now inspect the answer. If the client has additional requests, it can make them now. If it is done, it can use `DISCONNECT` to terminate the connection. Usually, an initial `DISCONNECT` is a blocking call, suspending the client and sending a packet to the server saying that the connection is no longer needed (5). When the server gets the packet, it also issues a `DISCONNECT` of its own, acknowledging the client and releasing the connection. When the server's packet (6) gets back to the client machine, the client process is released and the connection is broken. In a nutshell, this is how connection-oriented communication works.

Of course, life is not so simple. Many things can go wrong here. The timing can be wrong (e.g., the `CONNECT` is done before the `LISTEN`), packets can get lost, and much more. We will look at these issues in great detail later, but for the moment, above Figure briefly summarizes how client-server communication might work over a connection-oriented network.

Given that six packets are required to complete this protocol, one might wonder why a connectionless protocol is not used instead. The answer is that in a perfect world it could be, in which case only two packets would be needed: one for the request and one for the reply. However, in the face of large messages in either direction (e.g., a megabyte file), transmission errors, and lost packets, the situation changes. If the reply consisted of hundreds of packets, some of which could be lost during transmission, how would the client know if some pieces were missing? How would the client know whether the last packet actually received was really the last packet sent? Suppose that the client wanted a second file. How could it tell packet 1 from the second file from a lost packet 1 from the first file that suddenly found its way to the client? In short, in the real world, a simple request-reply protocol over an unreliable network is often inadequate.

3.5.5 The Relationship of Services to Protocols

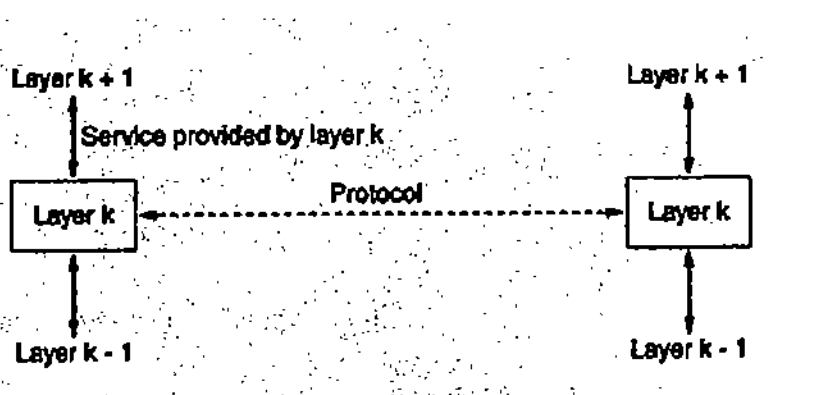
Services and protocols are distinct concepts, although they are frequently confused. This distinction is so important, however, that we emphasize it again here. A service is a set of primitives (operations) that a layer provides to the layer above it. The service defines what operations the layer is prepared to perform on behalf of its users, but it says nothing at all about how these operations are implemented. A service relates to an interface

between two layers, with the lower layer being the service provider and the upper layer being the service user.

A protocol, in contrast, is a set of rules governing the format and meaning of the packets, or messages that are exchanged by the peer entities within a layer. Entities use protocols to implement their service definitions. They are free to change their protocols at will, provided they do not change the service visible to their users. In this way, the service and the protocol are completely decoupled.

In other words, services relate to the interfaces between layers, as illustrated in figure. In contrast, protocols relate to the packets sent between peer entities on different machines. It is important not to confuse the two concepts.

Figure: The relationship between a service and a protocol



An analogy with programming languages is worth making. A service is like an abstract data type or an object in an object-oriented language. It defines operations that can be performed on an object but does not specify how these operations are implemented. A protocol relates to the implementation of the service and as such is not visible to the user of the service.

Many older protocols did not distinguish the service from the protocol. In effect, a typical layer might have had a service primitive SEND PACKET with the user providing a pointer to a fully assembled packet. This arrangement meant that all changes to the protocol were immediately visible to the users. Most network designers now regard such a design as a serious blunder.

3.6 Summary

In this unit, the concept of protocols and standards is elaborated. Networking protocols and their key elements are described. After that it discusses networking hardware requirements for different type of networks, for example, wireless networks, home networks and Internetworks. Finally network software is explained that includes protocol hierarchies, design issues, service primitives and the relationship of services to protocols.

3.7 Review Questions

1. What is a protocol? What are the elements of a networking protocol?
2. What is a standard? What are the key advantages of standardization?
3. What is the difference between broadcast link and point to point link?
4. Describe home networks.
5. Describe Internetworks.
6. What do you understand by protocol hierarchy?
7. What are the design issues for layers?
8. What is the difference between connectionless and connection-oriented services?
9. Describe service primitives.
10. Explain the relationship of services to protocols

Unit IV – Network Design And Configuration

What You Will Learn (Objectives)

After completing this unit you will be able to:

- Identify common resources shared within a network.
- Design and configure a network
- Question a network manager or client about network design, issues, and benefits.

4.1 Network Components/Configurations

In the lab, you are going to configure your computer, enabling it to share resources over a local area network. This involves enabling file and print share permissions and configuring client, protocol, and service components.

Windows XP CDs come with networking software. The software may already be installed on your computer workstation, or you may have to install it yourself from the Windows 9 CD. You may also need to download additional drivers from the Internet. See your instructor for assistance.

It is important to note that not all of the windows illustrated in this lab (and in subsequent labs) will look exactly like the windows on your personal desktop. If you have any questions regarding differences, see your instructor for assistance.

Materials Needed


- Networked Computer
 - Windows Operating System – CD-ROM
 - A folder and/or files to share
-

4.1.1 Directions



1. Work in teams of three or four.
2. As you progress through this lab, keep a record of your experiences in your log book. Include any observations you think important, such as the following:
 - Did you encounter any difficulties? What were they? How did you solve them?
 - How would you change the lab if you were to do it again?
 - What were some positive experiences?
 - How long did it take to complete the lab?
 - Did you understand what you were doing?
 - Any other comments you want to include.
3. Upon completion of the lab, discuss the group dynamics of teamwork.
 - Did your team work well together?

- Did you listen to your team members?
 - Did they listen to you?
4. After the teamwork discussion, write a short summary of your discussion and put your teamwork summary into your portfolio. Include observations such as the following:
 - Why do you think teamwork is important?
 - What was positive?
 - What was negative?
 - How might you improve interactions with team members in the future?

4.1.2 Procedure: Creating File and Print Sharing Permissions

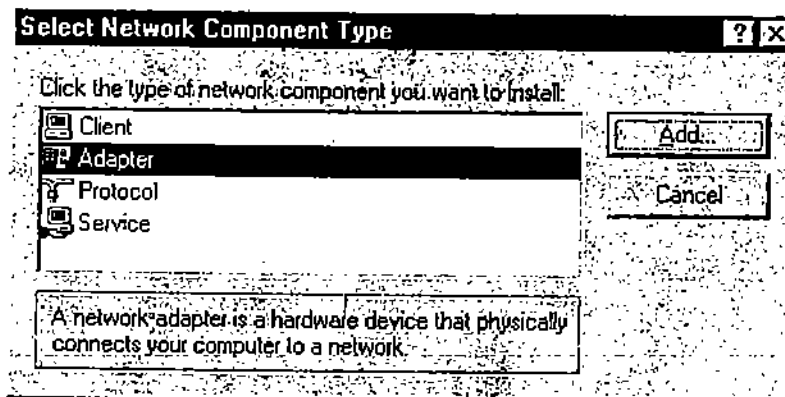
1. Start your computer.
2. Click 
 - Select Settings.
 - Select Control Panel.




3. In the Control Panel window, double-click the Network  icon.
4. In the Network window, highlight File and Print sharing for Microsoft Networks.
5. Click .

Configuring Adapters

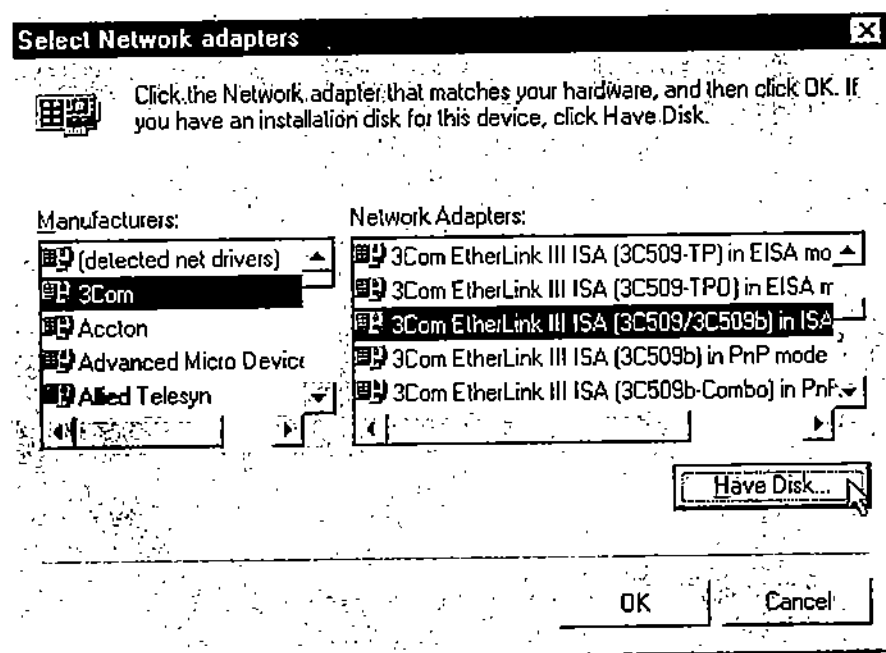
6. In the Select Network Component Type window, highlight Adapter.



7. Click .
8. A window similar to the one below will appear on the desktop.
 - See instructor for the name of the manufacturer of the network interface adapter installed in your computer workstation. If the manufacturer of your adapter does not appear in this window, you will have to download the driver

from the Internet or use the software that come with the adapter; see your instructor.

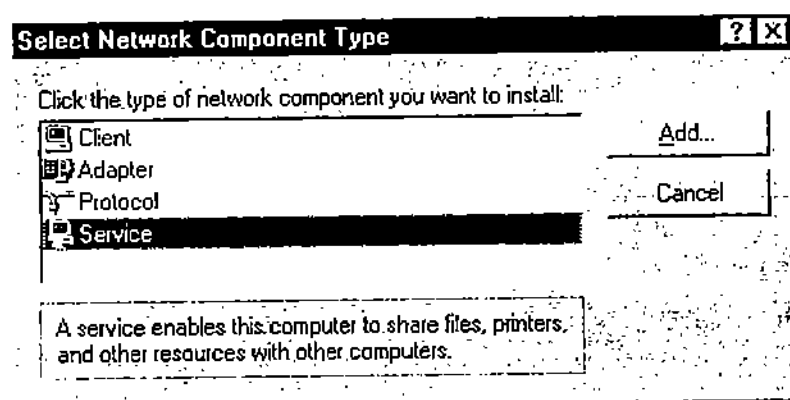
9. In the Select Network adapters window, highlight the correct manufacturer and network adapter.



10. Click OK.

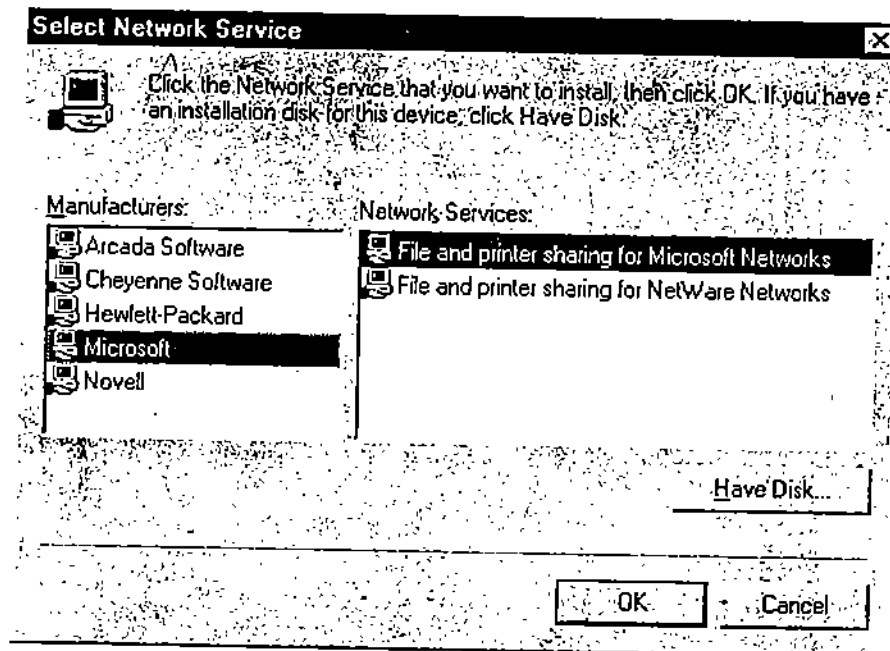
4.1.3 Configuring the Service Components

1. Follow steps 1-5 above under **Creating File and Print Sharing Permissions**.
2. In the select Network Component Type window, highlight Service.



3. Click Add....

4. In the Select Network Service window, highlight Microsoft and File printer sharing for Microsoft Networks. (At this point, you may have to install the network service. See your instructor.)

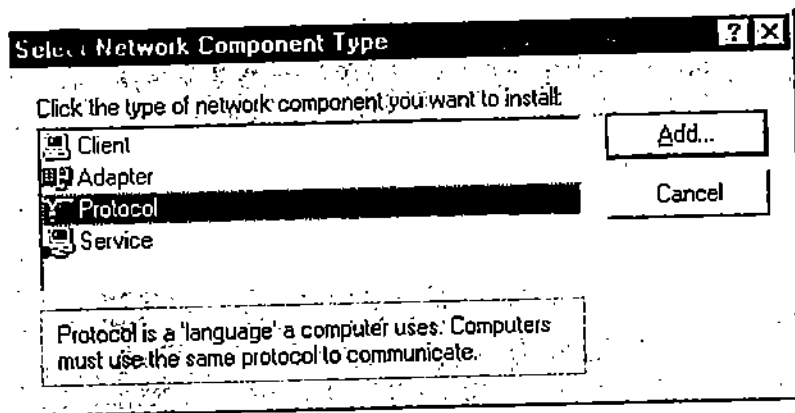



5. Click .
- 6.

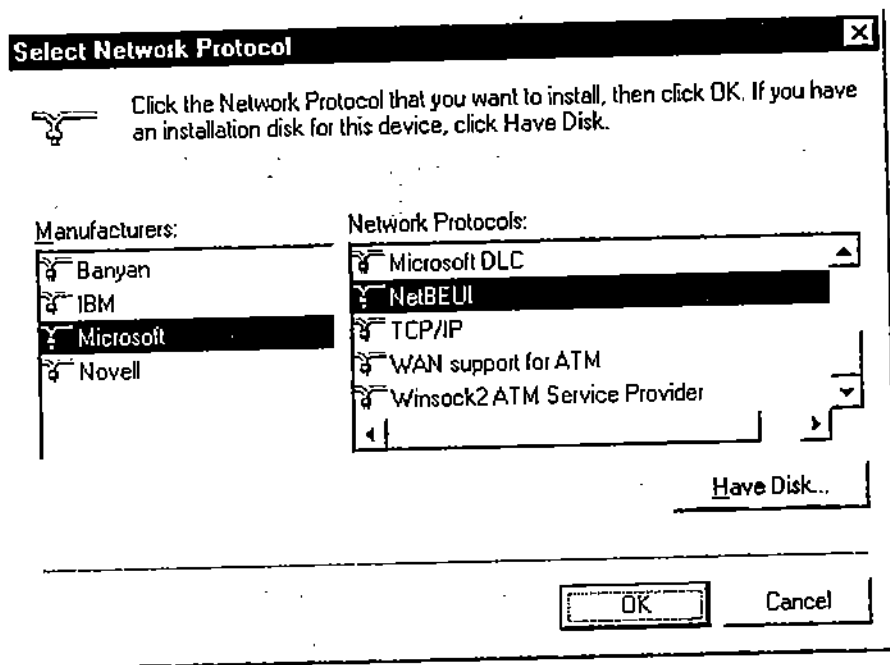
4.1.4 Configuring Protocol Components




NetBEUI is a protocol used in small LANs for communications over a network. You will learn more about NetBEUI later in this course. It comes with Windows 95/98. Before enabling NetBEUI to communicate over a network, its driver must be installed on your computer. In this activity, you will set up the NetBEUI protocols for your workstation.

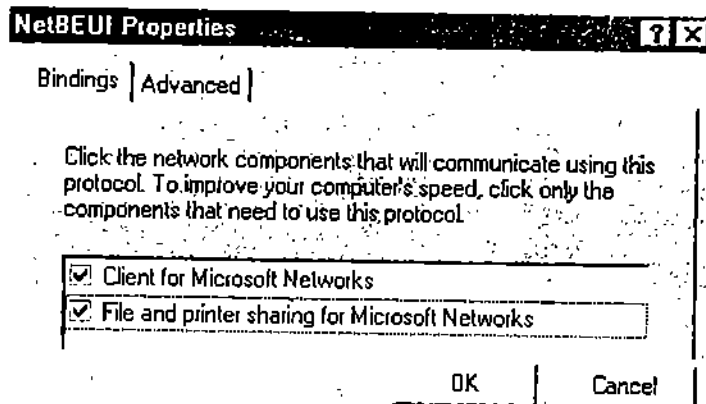
1. Follow steps 1-5 above under **Creating File and Print Sharing Permissions**.
2. In the select Network Component Type window, highlight Protocol.



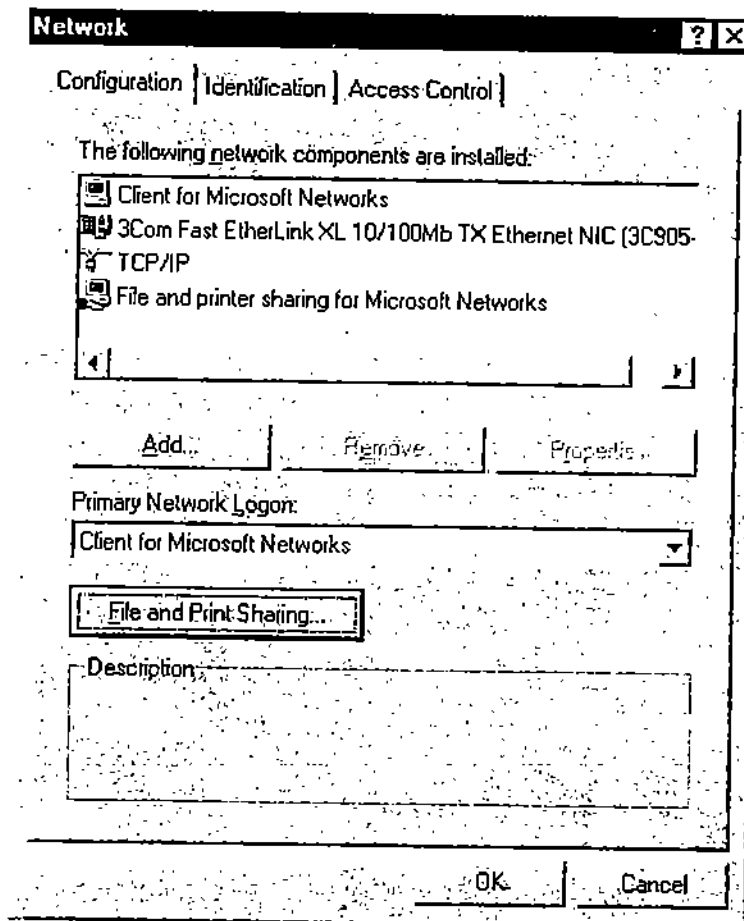
3. Click .
4. In the Select Network Service window, highlight Microsoft and NetBEUI.



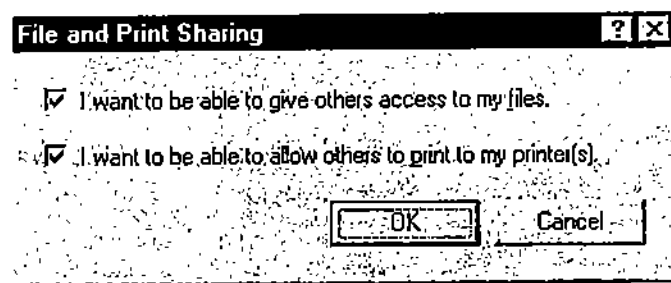
5. Click .
 6. You will see Network Configuration window next. Notice that the NetBEUI protocol has been added.
 7. In the Network Configuration window, highlight NetBEUI (manufacturer name).
 8. Click on  and check to see that Client for Microsoft Networks and File and Printer sharing for Microsoft Network boxes are checked. If they are not checked, please select them and click .
- Reminder:** your window may not look the same as the illustration below.

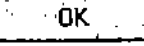


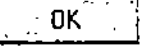
9. At this point, you are ready to share the folder you created at the beginning of this lab.
10. On the Configuration tab page in the Network window, click File and Print Sharing...



11. Select both the printer and file sharing boxes.



12. Click . This will return you to the Configuration tab page of the Network window.

13. Click  to close the Network dialog box.

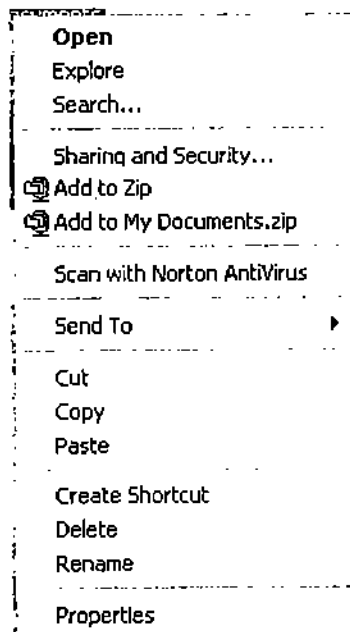


14. Double-click the  icon on the desktop.

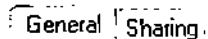
15. Double-click the **(C:)** Drive icon.

16. Check with your teacher for the name of the folder and or files to be shared.
Right-click this folder.

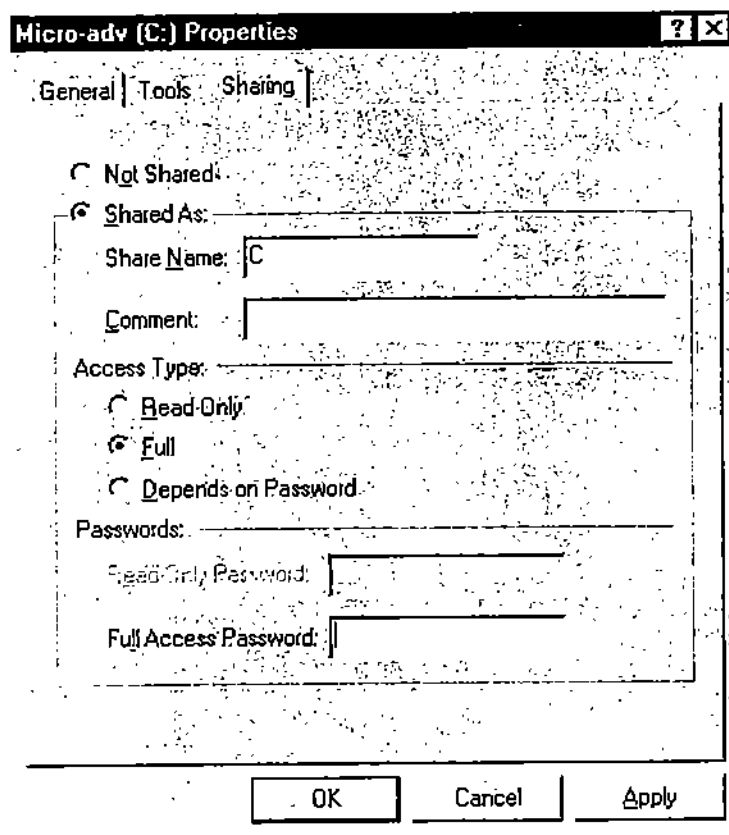
17. Highlight Properties.



18. In the "Name of the Document" window, click the **Sharing** tab.



19. Click the Shared As button and the Full button.



20. When you click Full, the folder and or files can be accessed, read, and modified by anyone on the network.


- Why would you want to give a user full access to a folder?
- Describe a situation where full access might be needed.
- Why would you want to give read-only access to a user?
- Describe a situation where read-only access might be important.
- Why would you want to use a password?
- Describe a situation where you would want to use a password for user access.

21. Click Apply.


22. Click OK.

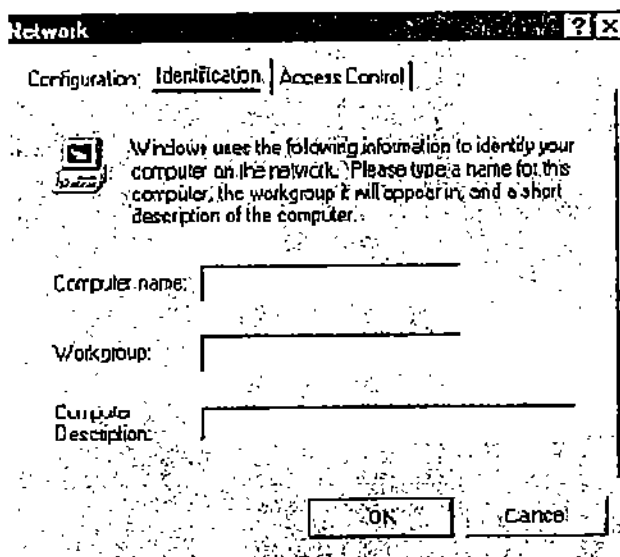
- What happened to the folder and or files you are sharing?
- Who can access this folder and files? Why?
- What is the significance of the hand on the folder and or files?
- What are some resources that you can share over a network?

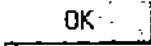
4.1.5 Procedures: Identifying Your Computer for a LAN

1. Start your computer.
2. Click .
 - Select Settings.
 - Select Control Panel.



3. In the Control Panel window, double-click the Network  icon.
4. In the **Network** window, click the **Identification** tab.



1. Enter the computer name and workgroup name supplied by your instructor.
2. Click .

4.1.6 Procedures: Connect to a LAN and share Resources

Earlier in this lab, you created a folder and files to share. You then configured your computer, enabling it to communicate over a network. Now you will access a shared resource from one of your peers.

1. From the desktop window, double-click on **Network Neighborhood**.
2. Double-click on **Entire Network**.
3. Computers in your workgroup will appear as icons.
4. Double-click your workgroup, folder and/or files to be shared.
5. At this point, follow your instructor's directions.

4.2 Interview and Role-Play

Directions

1. Pair with another student in the class for this activity.
 2. Together, in person or by phone, interview the manager of a supermarket chain or another company that uses a network.
 3. Prior to the interview, create a questionnaire to use as a checklist and as a place to record your notes or answers. Avoid questions with "yes" or "no" answers. When you have finalized your questionnaire, place a copy in your portfolio for future use.
 4. In your questions to the manager, inquire about the advantages and disadvantages of the network system and what changes would he or she make to improve the system.
 5. Immediately after the interview, expand your notes with details of your interviewee's answers. Record the responses in your portfolio.
 6. With the questionnaire and responses in hand, create a script for role-playing your interview in front of the class. One of you will be the interviewer and the other the interviewee.
 7. Remember that you must demonstrate the same professionalism in the role-play that you did during the actual interview. Be creative in your presentation to the class. Make it interesting.
-

4.2.1 Procedure: Conducting the Interview-Sample Scripts & Questions

Sample Script for Interview

- To the receptionist or whomever answers the phone:
 1. Hello. Our names are (your name) and (your partner's name).
 2. We are students in (teacher's name) Internetworking Fundamentals course at (school's name).
 3. Our assignment is to interview a manager of a networked company's network.
 4. May we speak to the manager of your company for a few minutes?
 5. We have just a few questions to ask about the operation of your network.
- To the company manager:
 1. (Repeat numbers 1-3 from above).
 2. Is now a good time to answer our questions? We estimate needing 15 minutes of your time. Or would another time be convenient?
 3. Thank the manager for his/her time and either proceed with the interview or set another time.

Sample questions

1. What type of network do you have?
2. What type of changes have you seen take place in your company due to the addition of networking technology?
3. Has the networking technology been difficult for the employees to learn?
4. What problems have you or your employees had using the network?
5. How does your company use the network?
6. Has the network increased productivity or collection in your company?
7. What resources are shared over your network?
8. What do you see as the major advantages and disadvantages of using a network in your company?
9. What changes would you make to the network?
10. Do you have a designated network administrator? What is the position title? Is the person full or part-time?
11. How many people maintain the network?
12. Do you subcontract for any services?

Checklist for the Interview

1. Be sure to have your questions prepared ahead of time. You do not want to take more time than necessary.
2. If you go in person, consider your appearance. You represent not only your school, but also yourself. One day you may be designing a network for this company.
3. Take a pen or pencil and writing paper. You will want to appear efficient.
4. Be polite and be sure to thank your interviewee.
5. Follow up your interview with a short thank you note by mail or email.

4.3 Network Wizards

Researching a Network

Directions

1. Select and research how networking is used in one of the following environments:
 - Airline/hotel/concert reservations
 - Telephone
 - Cable television
 - The Internet
2. Use the supermarket example found in the content manual of this lesson as a guide.
3. Diagram the path of a transaction across the network.
4. Present your diagram and research in an oral presentation to the class.
5. Prepare a list of the resources and references used to complete to your portfolio.
6. Add your diagram, resource list, and presentation outline to your portfolio.

4.4 Summary

This unit describes the procedure of configuring and designing a network that also includes the network components. We learn to identify the common resources shared within a network. Finally it explains that how to develop a questionnaire and interview a network manager or a client about network design, issues, and benefits.

4.5 Review Questions

Network Design Questions

1. Draw a diagram of a Local Area Network. Label and show both a Peer-to-Peer and Client/Server environment.
2. Draw a diagram of an Enterprise Network. Label and include a LAN and a MAN in your diagram.
3. Draw a diagram of a Metropolitan Area Network. Label the networks.

Part C

Short Essay

Describe Peer-to-Peer networking and Client/Server networking.

Part D

List at least 5 common resources shared over a network.



Uttar Pradesh
Rajarshi Tandon Open University

BLISS-03

**Computer Network and
Security Maintenance**

BLOCK

2

Transmission and Network Elements

Unit 1 Signal Transmission	5
Unit 2 OSI Reference Model	18
Unit 3 Ethernet	34
Unit 4 Networks and Devices	48

Block Introduction:

This is the second block of the Data Communication and Networking. In this block signal transmission, OSI Reference Model, Ethernet, and different networks and devices are introduced.

In unit 1, there is introduction to different types of transmission media and signaling. The packet and frame structure and their relation to OSI Model are also explained.

Unit 2 starts by stating the importance of OSI Model. It lists all the seven layers, identify the functions performed by each layer, and briefly explains that how they communicate with each other.

In unit 3 Ethernet is introduced. It describes Ethernet topology and how CSMA/CD is used within an Ethernet LAN. Finally, it covers the relationship between IEEE Ethernet standards and the OSI Model.

Unit 4 is mainly focused about networks and devices. It describes ATM, Token ring and FDDI architecture. It also covers different connectivity and internetworking devices with their functionality.

Unit I – Signal Transmission

What You Will Learn (Objectives)

After completing this unit you will be able to:

- Define technical terms associated with data signaling and transmission.
- Learn different types of transmission media
- Describe the characteristics of digital and analog signaling.
- Explain how packets and frames are structured, and describe their relationship to the OSI model.

Terminology

- **Bandwidth**~ Bandwidth represents how much information can be sent at one time over a specific cable type, or the measure of information capacity of transmission over a cable. Bandwidth is the difference between the highest and lowest frequencies of the transmission, measured in hertz. Bandwidth is a factor used to evaluate cable.
- **BNC**~ The "British Naval Connector" is the connector used with coaxial cables.
- **Fiber Optic Cable**~ Fiber optic cable uses light to transmit information across a network. The core of the cable is made of glass, which is protected by a layer of gel or plastic. A plastic cover surrounds the entire cable.
- **EMI (Electro-Magnetic Interference)**~The interference of electrical signals across a cable by outside electrical or magnetic devices. It is a factor that is used to evaluate cable.
- **Transmission Media**~ Transmission media is the physical pathway used to connect two or more computers together as a network and includes air and space in addition to cable.
 - **Amplitude Characteristic** of a wave measuring wave height from the base to the peak of a waveform indicates the strength of the signal.
 - **Analog Signal**: Analog signals change continuously as opposed to digital signals, which are discretely valued. For example, sound is an analog signal; it is continuous and varies in strength.
 - **ASCII Code**: American Standard Code for Information Interchange. A 7-bit coding scheme that assigns unique numeric values to letters, numbers, punctuation, and control characters.
 - **Baudot Code**: A 5-bit coding scheme used for transmitting data.
 - **Binary Numbers**: A number system based on two states, 0 and 1. Computers use combinations of binary numbers to represent and encode all kinds of data including words, sounds, colors, and pictures.
 - **Connection-Oriented Communication**: A form of network communication, where the transmitting device must establish a connection with the receiving

device before data can be transmitted, (for example, telephone). In connection-oriented communication, the receiving device acknowledges receipt of the data.

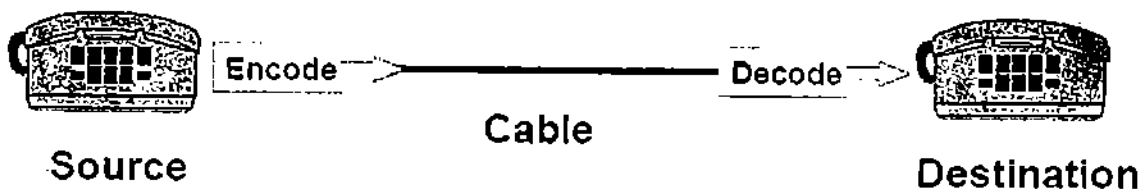
- **Connectionless Communication:** A form of communication over networks where the transmitting device can send a message without establishing a connection with the receiving device (for example, radio). Signals are sent, but there is no mechanism for acknowledging receipt.
- **Digital Signal:** Data transmitted in discrete states, for example, on and off. These discrete states can be represented by binary numbers, and vice versa.
- **Full-Duplex:** Two-way, simultaneous data transmission. Each device has a separate communication channel.
- **EBCDIC Code:** Extended Binary Coded Decimal Interchange Code. An 8-bit coding scheme used by IBM for data representation in mainframe environments.
- **Logical Address:** An OSI model Layer 3 address.
- **Frame:** Basic unit of data transfer at OSI Layer 2.
- **Half-Duplex:** Two-way data transmission that is not simultaneous. Only one device can communicate at a time.
- **Packet:** Basic unit of data transfer at OSI Layer 3.
- **Physical Address:** A OSI model Layer 2 address.

1.1 Data Transmission

Data is transmitted over networks using signals, which are transformed, or encoded, by computers into the voice, video, graphics, and/or the print we see on our computer screens. The signals used by computers to transmit data are either digital or analog.

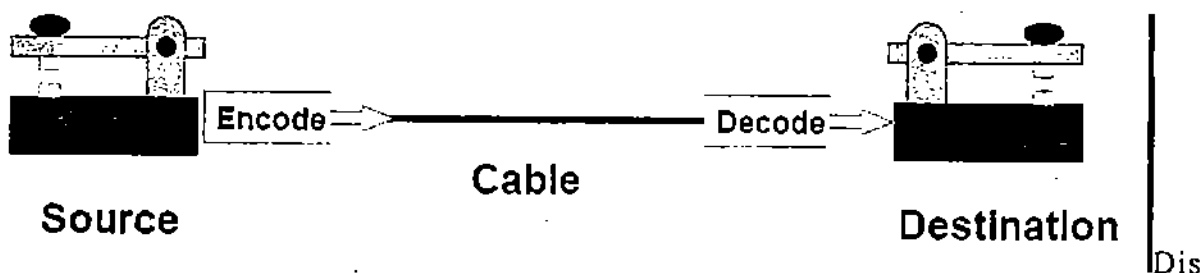
Analog signals are continuous signals that vary in strength. Sound is an example of an analog signal. Sound is actually a wave and is quite similar, or analogous, to electromagnetic waves, hence the name analog. Telephones have transmitters that encode sound waves into electromagnetic waves, which then travel over wires toward their destination. The receiving telephone decodes the electromagnetic waves back into sound waves. Our brains then decode the sound waves into the words we hear. Computer modems use the same principle. Analog signals can be represented digitally. For instance, a high electromagnetic voltage could be interpreted as 1 and low voltage as 0.

Telephone Encoding/Decoding



Digital signals are discrete rather than continuous. Either there is a signal or there isn't a signal. Telegraphs transmit data with discrete signals. You either hear a tap or you do not hear a tap. Discrete signals can be represented by on and off pulses. The duration of a discrete signal can be varied, as with dots and dashes in Morse code.

Telegraph Encoding/Decoding



crete signals can also be represented digitally. The presence of a signal could be coded as a 1 and the absence of a signal coded as a 0. The digits 0 and 1 are used because computer circuitry is based on binary digital data. Codes are used to group a set number of bits together and have a group of bits represent a letter, number, or other character. The computer's brain, the central processing unit (CPU), transforms these codes of 0s and 1s into the voice, video and data we see. One coding scheme, ASCII, codes an "a" as the binary number 01100001.

Digital data is based on two states, on or off. The binary numbering system uses only two digits, 0 and 1, so it makes sense to use the binary numbering system. One digit, 0 represents off, the other digit represents on. A single 0 or 1 is called a bit. One byte is equal to eight bits (also called an octet when discussing TCP/IP). In ASCII code, one octet is the equivalent of one alphabetic or numeric character. In order to appreciate how computers communicate over networks, it is necessary to be aware of how they encode information.

1.2 Connection-oriented and Connectionless Transmissions

Data transmission may be:

- Connection-oriented

- **Connectionless**

The main difference between the two is that with a connection-oriented transmission, the destination device acknowledges receipt. Whereas, with connectionless, there is no acknowledgement.

In connection-oriented transmissions, the sending (source) device establishes a connection with the receiving (destination) device. The connection is continued until all data packets have been transmitted and the source device receives notification that the data was received by the destination device and has been checked for errors.

A telephone conversation is an example of a connection-oriented transmission. When a call is made, data is transmitted across phone lines, the receiving party picks up the phone, and a conversation takes place. The individual making the call knows that it arrived at the correct destination and that it was understood.

In a connectionless transmission, the source device transmits data but the connection is not maintained. The source device does not wait for notification that the destination device actually received the information accurately. This method is faster than connection-oriented, however less reliable since there is no notification of whether the data is received or not. It is more common to find connectionless transmissions on LANs.

To understand a connectionless transmission, think of a radio broadcast: A radio disc jockey tells his/her friends to be sure to listen to her/his program at 9:00 p.m. At that time the disk jockey broadcasts a message to them. Did they receive the message? Although it is quite likely, the disk jockey cannot be sure that they turned the radio on, listened, or understood the message.

1.3 Synchronous and Asynchronous Transmission

Computers need to know when to expect data and where a character begins and ends. When receiving data, timing on both computer devices must be coordinated if they are to work together efficiently. This coordination is called clocking, timing, or framing. There are two protocols for the timing or coordination of data signals:

- Synchronous
- Asynchronous

When transferring data, both the transmitting and receiving nodes need to agree when the signal begins and ends so the signals can be correctly measured and interpreted. This timing process is called bit synchronization, framing, or clocking.

Imagine how difficult it would be to read if you did not know when a word started and when a word ended if there were no punctuation and no spaces you can do it because there are several different characters and it is not in code what if this were coded as zeros and ones then you would have real problems. As you can see, synchronization of data is very important.

Clocking is somewhat like timing in music. There are a specific number of beats expected per bar. When computer devices are synchronized, a specific number of signals or "beats" are expected within a set amount of time. Timing is important because it helps you be prepared. In many schools, every 50 minutes, a new class

period starts. Students watch the clock and expect a signal. Usually, they are already prepared to leave the classroom. That is because they expected the signal.

Synchronous transmission requires the communicating devices to maintain synchronous clocks during the entire connection. The sending device transmits on a specific schedule and the receiving device accepts the data on that same fixed schedule. The receiving device knows the timing of the sending device because the timing information is embedded within the preamble of the frame. Synchronous transmissions are common in internal computer communications and usually are sent as entire frames. Synchronous transmission is common when large blocks of data are transferred, since it is efficient and has a low overhead (number of bytes of data/control + data bytes).

Asynchronous data transmission does not involve synchronizing the clocks of the sending and receiving devices. Instead, start and stop bits are used for synchronization of data signals. The start and stop bits tell the receiving device how to interpret the data. Asynchronous sends one character at a time.

Data transmission may be half-duplex; meaning data is transferred in only one direction at a time. An example of half-duplex is a CB radio where only one person can talk at a time. Or, transmission may be full duplex, transmitted in two directions simultaneously. A telephone conversation illustrates full-duplex communication.

1.4 Transmission Media

A **transmission medium** (plural *transmission media*) is a material substance (solid, liquid or gas) which can propagate energy waves. For example, the transmission medium for sound received by the ears is usually air, but solids and liquids may also act as transmission media for sound.

The absence of a material medium (the vacuum or empty space) can also be thought of as a transmission medium for electromagnetic waves such as light and radio waves. While material substance is not required for electromagnetic waves to propagate, such waves are usually affected by the transmission media through which they pass, for instance by absorption or by reflection or refraction at the interfaces between media.

The term **transmission medium** can also refer to the technical device which employs the material substance to transmit or guide the waves. Thus an optical fiber or a copper cable can be referred to as a transmission medium.

Electromagnetic radiation can be transmitted through an optical media, such as optical fiber, or through twisted pair wires, coaxial cable, or dielectric-slab waveguides. It may also pass through any physical material which is transparent to the specific wavelength, such as water, air, glass, or concrete. Sound is, by definition, the vibration of matter, so it requires a physical medium for transmission, as does other kinds of mechanical waves and heat energy. Historically, various other theories were used in science and thought to be necessary to explain the transmission medium. However, it is now known that electromagnetic waves do not require a physical transmission medium, and so can travel through the "vacuum" of free space. Regions of the insulative vacuum can become conductive for electrical conduction through the presence of free electrons, holes, or ions.

Telecommunications

Many transmission media are used as a communications channel.

For telecommunications purposes in the United States, Federal Standard 1037C, transmission media are classified as one of the following:

- Guided (or bounded) - Waves are guided along a solid medium such as a transmission line.
- Wireless (or unguided) - Transmission and reception are achieved by means of an antenna.

Wireless media may carry surface waves or skywaves, either longitudinally or transversely, and are so classified.

1.4.1 Decision Factors in Choosing Cable For a network

- The *bandwidth* represents how much information can be transferred over the cable at one time. The greater the cable's bandwidth the faster the information is passed over the network.
- The further information must travel over a cable the more the electrical strength weakens. This is referred to as *signal attenuation*. Each type of cable can transmit information up to a specific distance without special equipment to help boost the electrical signal.
- *EMI (Electro-Magnetic Interference)* comes from electrical signals emitted by equipment such as photocopiers, fluorescent lights, and electrical wiring that may disrupt the transmission of information across a network. Some cables are better designed to block out EMI (Electro-Magnetic Interference) than others are.
- Organizations may not always be able to build the largest network needed, or their office may grow over time and need more computers added to their network. The ease of installation and need for additional equipment affect the *expansion capability* of the cable.
- The *cost* of the cable used to build a network is very important. Most organizations have limited funds and seek to build the optimum network.

CABLES

Solid wire and cable are the oldest forms of electronic transmission media. This lesson covers three basic types, still in use in building networks, coaxial, unshielded twisted-pair, and fiber optic. Thin coaxial cable has a core of copper wire and is primarily used for peer-to-peer LANs due to its low bandwidth and problems with EMI (Electro-Magnetic Interference). Unshielded twisted-pair cable has twisted pairs of wires as the core and is divided into five categories, with category 5 used most commonly for building LANs. Fiber optic cable has a core made of glass and uses light pulses to transmit information across a network.

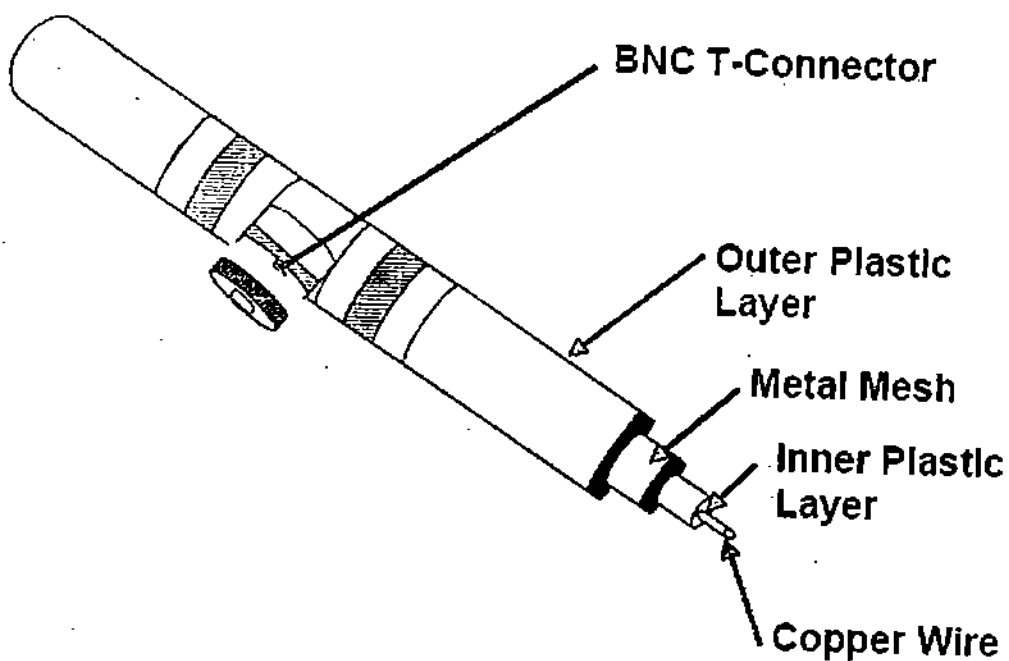
Thin Coaxial Cable

Early networks used coaxial cable to connect computers together. Many LANs were built with coaxial cable. It is often referred to as *ThinNet*.

Coaxial cable has:

- A core of copper wire surrounded by a layer of plastic.
- A layer of metal mesh.
- An outer protective plastic insulation sheath.

Thin Coaxial Cable



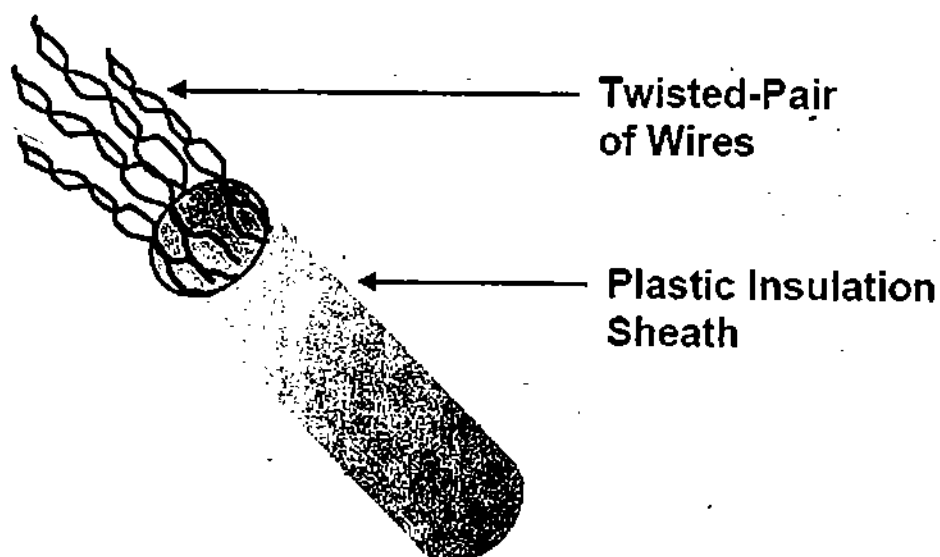
Decision Factor	Thin Coaxial Cable
Maximum Bandwidth	10 megabits per second.
EMI (Electro-Magnetic Interference)	Significant problems with neighboring electrical equipment.
Signal Attenuation	Maximum distance is 185 meters.
Expansion Issues	Expansion into multiple rooms is difficult since each computer must be connected directly on to the cable in a chain fashion, often referred to as a <i>bus topology</i> . The cable is thicker and less flexible than unshielded twisted-pair cable. This cable is best used for peer-to-peer networking in a small workgroup LAN.
Relative Cost	Low.

Unshielded Twisted-pair Cable

Unshielded twisted-pair cable is separated into five categories designated by the TIA/EIA 568-A standard.

- Category 1 is telephone cable.
- Category 2 was used for token ring networks and is not recommended for Ethernet networks.
- Categories 3 and 4 can be used with Ethernet networks, but suffer more from EMI than category 5. Category 3 cables typically have two twists per foot. Category 4 cables have more twists per foot, but less than Category 5 cables. The twisting of the wires in cables is to help prevent EMI (Electro-Magnetic Interference).
- Category 5 cable is primarily used in LANs. The most typical connector used with UTP is a RJ-45, which resembles a large telephone connector (RJ-11).

Unshielded Twisted-Pair Cable



Decision Factor	Unshielded Twisted-Pair Cable: Category 5
Maximum Bandwidth	100 megabits per second.
EMI (Electro-Magnetic Interference)	Each pair is twisted around each other at a different rate, which reduces EMI (Electro-Magnetic Interference) between the pairs and other electrical signals.
Signal Attenuation	Maximum distance is 100 meters.
Expansion Issues	<p>The cable is thin and flexible, which makes installation easy. Expansion is easy, but specialized network equipment is needed to boost the signal. This equipment increases the cost of expansion.</p> <p>This cable is used as a standard today in all Ethernet LANs.</p>
Relative Cost	Least expensive.

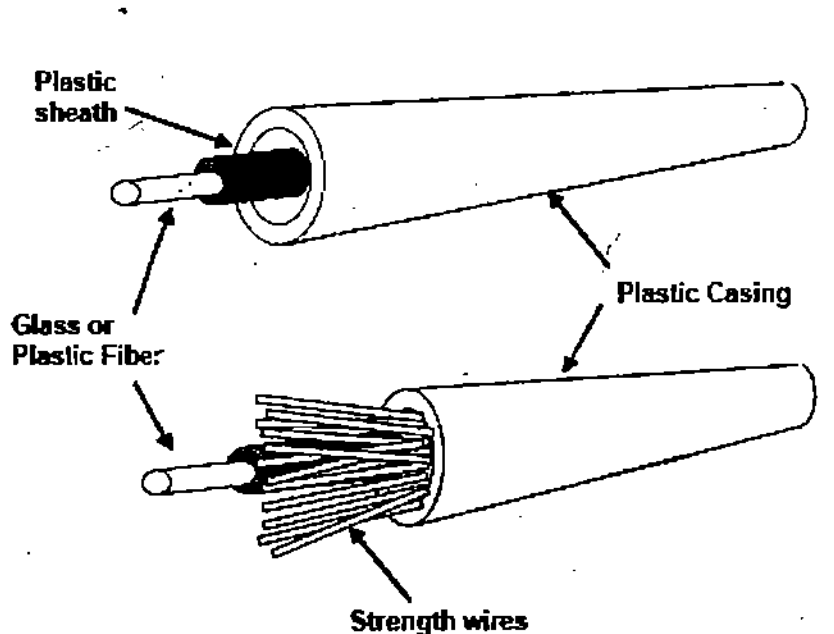
Within UTP cable, there are typically 2 to 8 wires, 1 to 4 pairs. Each pair is twisted around each other at a different rate, which reduces EMI (Electro-Magnetic Interference) between the pairs and other electrical signals. Each pair is color-coded, according to standards, and the entire cable is inside a protective plastic insulation sheath.

In Category 5 UTP, there are usually 4 pairs of wires, with one wire of solid color and one of white with the same color stripe in each pair:

- 1 orange pair
- 1 brown pair
- 1 blue pair
- 1 green pair

Fiber Optic Cable

Fiber optic cable uses light pulses rather than electrical signals to transmit information across a network. The cable may be used over many miles because there is no electrical EMI (Electro-Magnetic Interference) and the bandwidth is very high. Fiber optic cable is usually used for the backbone of a network. Since glass and plastic cores can be cracked or broken, installation requires care. Special monitoring equipment is required to locate a break in the fiber optic cable.



The core cable may be made of glass or plastic, which is protected by layer of

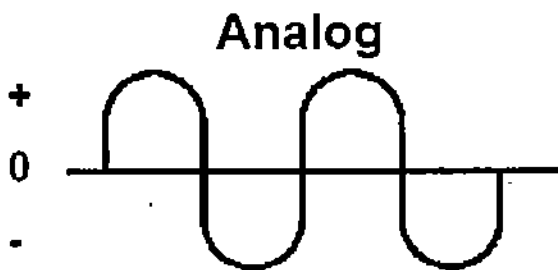
gel or light reflecting plastic. A plastic insulation sheath then surrounds the entire cable. Fiber Optic transmission speeds of 10Gigabits per second, currently, 1.6 Tera bits per second is being tested.

Decision Factor	Fiber Optic Cable
Maximum Bandwidth	10 Gigabits per second currently, though 1.6 Terabits per second is being tested.
EMI (Electro-Magnetic Interference)	No EMI (Electro-Magnetic Interference) occurs between the cable and other equipment.
Signal Attenuation	This cable may be laid over several miles.
Expansion Issues	<p>The use of fiber optic cable is typically limited due to the higher cost and difficulty in installation and maintenance. Fiber optic cable is easily broken requiring monitoring equipment to locate a break in the cable.</p> <p>The cable may be used as the backbone or the main cable of LANs, WANs, and MANs in Enterprise networks.</p>
Relative Cost	Most expensive.

1.5 Analog Signals

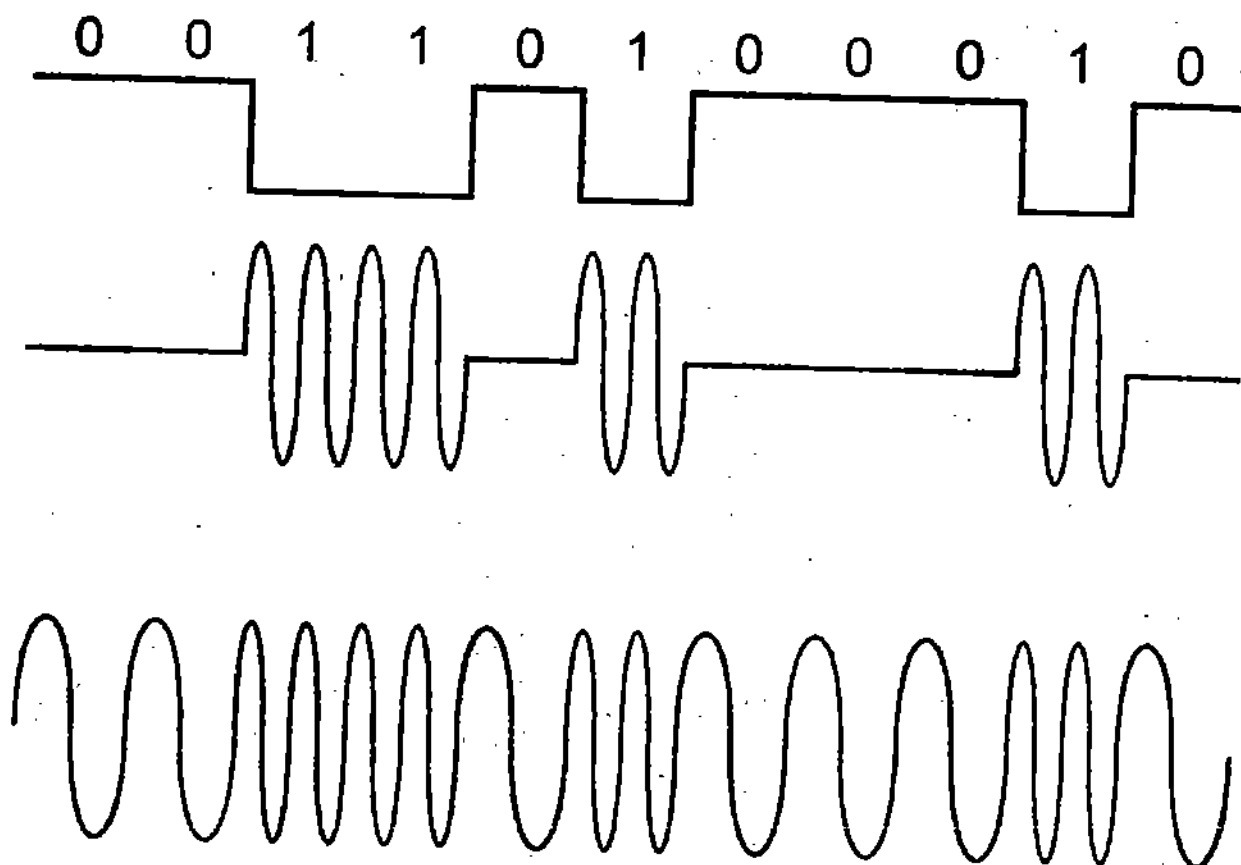
Analog signals, which are electromagnetic waves, are continuous and look like a copy of the original sound wave. Transmission of data is accomplished by varying one or more the waves' properties.

Analog Signal



All waves have three characteristics, amplitude (strength), frequency, and phase. Variations, called modulations, in wave characteristics are used to encode analog signals to digital signals. Amplitude-Shift Keying (variations in strength) and Frequency-Shift Keying are two examples. Amplitude-Shift Keying uses a change of the voltages for interpretation. When there is a voltage change from high to low, the binary digit represented changes. If high voltage were 1 then low voltage would be 0.

Amplitude Shift Keying



Frequency Shift Keying

Frequency-Shift Keying uses the frequency of the waves for interpretation. When there is a frequency change from high to low, the binary digit changes. If high frequency were 1 then low frequency would be 0.

Unit II – OSI Reference Model

What You Will Learn (Objectives)

After completing this unit you will be able to:

- Explain the importance of the OSI model.
- List the seven layers of the OSI model and describe how they communicate with one another.
- Identify the functions performed by each of the OSI model layers.
- Match network elements with their appropriate OSI model layer.
- Demonstrate an understanding of the OSI model in the context of information technology careers.

Terminology

- **Host**~ A host is a computer system (or node) that shares its resources within a network.
- **International Standardization Organization (ISO)**~The ISO is a non-governmental group of standards-setting organizations from more than 90 countries. The ISO proposes, revises, and publishes global standards and protocols for a wide variety of concerns including banking services, electrical products, sports equipment, toy safety, and so on.
- **Media** ~The physical material used when transmitting information over a network is referred to as the media. This includes conductive cables, optical fiber, or air through which encoded information flows between sender(s) and receiver(s), routers, switches, and so on.
- **Network Architecture**~ The overall connection design and layout for how computers, devices, and media transfer information on a network is described as the network architecture. The three main types of network architecture are star (ARCnet), Ethernet, and Token Ring.
- **Open Systems**~ Networks working together for efficient information flow by jointly implementing protocols whose details are publicly available are called open systems. Open systems contrast with proprietary systems.
- **Open Systems Interconnection (OSI) Reference Model**~The OSI model is a 7-layered format for organizing and describing the standardization required to enable full communication between different computers and computer systems.

- Operating System~ An operating system is software used to control the activity of an individual workstation and/or network. Examples include: DOS, Windows, and Macintosh OS for workstations and NetWare, UNIX, LINUX, Windows 95, and NT for networks.

- Peer Entities~ Parties or resources involved in communication at the same level are referred to as peer entities.
- Proprietary Systems~ Vendor-specific sets of protocols that are not typically supported by other vendors. Token Ring started as a proprietary system.
- Protocol~ Rules that govern network data flow and product creation. Protocols are pre-established communication rules and conventions.

1.1 Importance and Impact of the OSI Model

The OSI model is important because it:

- Describes a non-proprietary, open system of interfaces and protocols for network interconnection.
- Provides a framework for standards development for productive cooperation among global competitive vendors.
- Can be generalized as a model across electronic, electromagnetic, and optical communications technologies.
- Encourages design for information sharing regardless of computer/network architecture, topology, or platform.
- Balances protection against obsolescence with opportunities for improvement.
- Assists systematic understanding of complex network design and troubleshooting.

Protocols Govern Network Communication

Open systems networks are based on agreed-upon rules that allow users to communicate with each other. These communication rules are called *protocols* and are used by product developers and system designers to create computer networks. Due to the number of protocols needed to make different equipment communicate over a variety of media, networking is a complex task. Standards and models have been developed to describe the topics covered by networking protocols from a general or higher level perspective.

1.2 The OSI Model

The OSI model uses the concept of seven stacked layers to define a network communications system.

The lower three layers, listed below, deal mostly with network-dependent (hardware) functions. Transmissions of bits, frames, and packets over LANs, MANs, and WANs are handled by these layers.

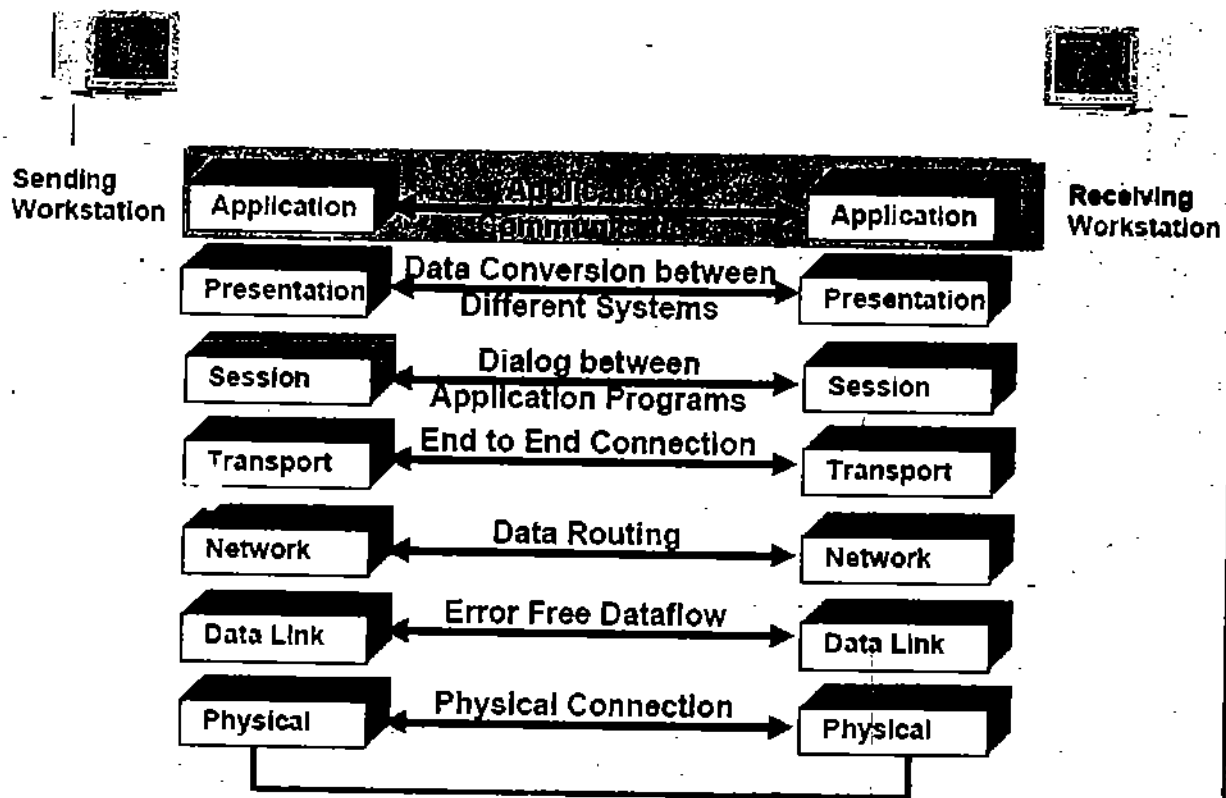
- Physical
- Data Link
- Network

The upper four layers, listed below, deal more with application-oriented (user or software) functions. Network operating systems applications such as Windows NT, Novell NetWare, and UNIX function at these upper layers.

- Transport
- Session
- Presentation
- Application

Each layer of the OSI model provides specific functions not provided by any other layer. Sets of protocols allow the peer layers of the sending and receiving computers to communicate with each other. Each level uses the services of the layer immediately below it and provides services to the layer immediately above it.

The OSI Model



1.3 Seven Layers of the OSI Model

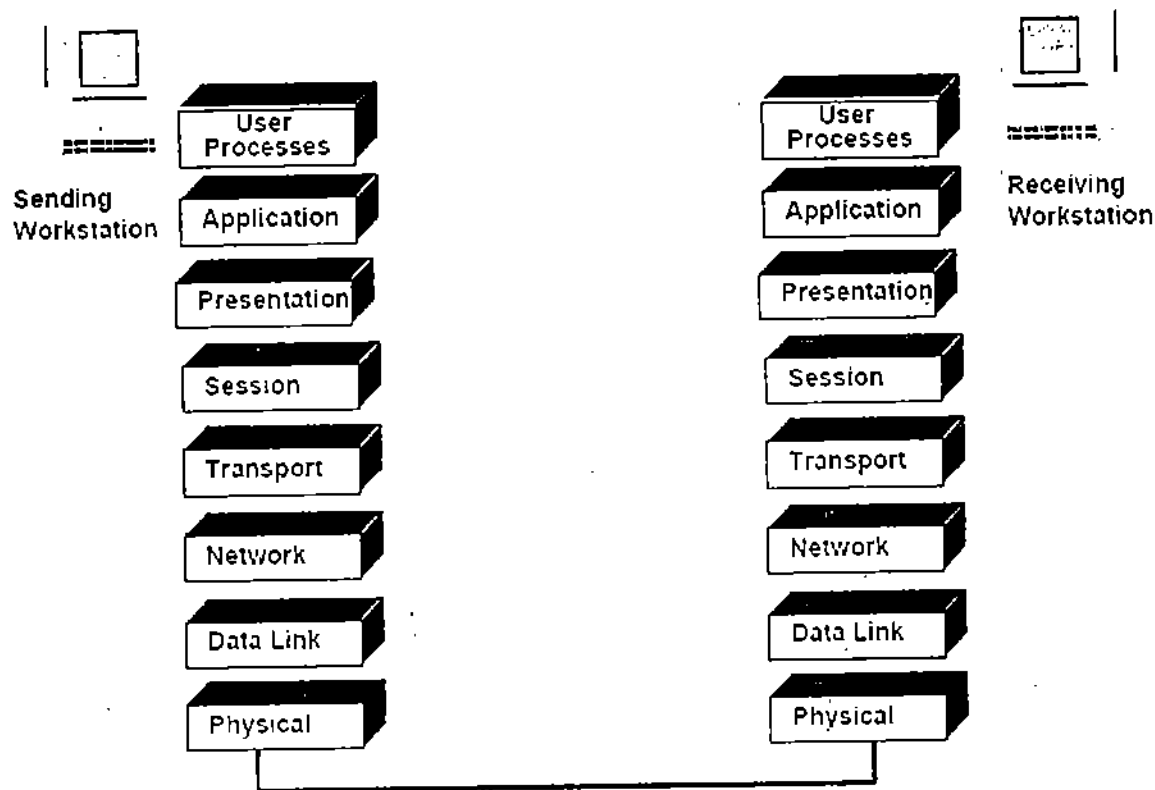
Layer 1: Physical

The OSI physical layer protocols handle the mechanical, electrical, and procedural functions needed to provide the physical link between devices and the network. An easy way to think about it is that the physical layer is anything you can touch that transmits the data—a plug, cable, fiber, and so on.

The Physical Layer

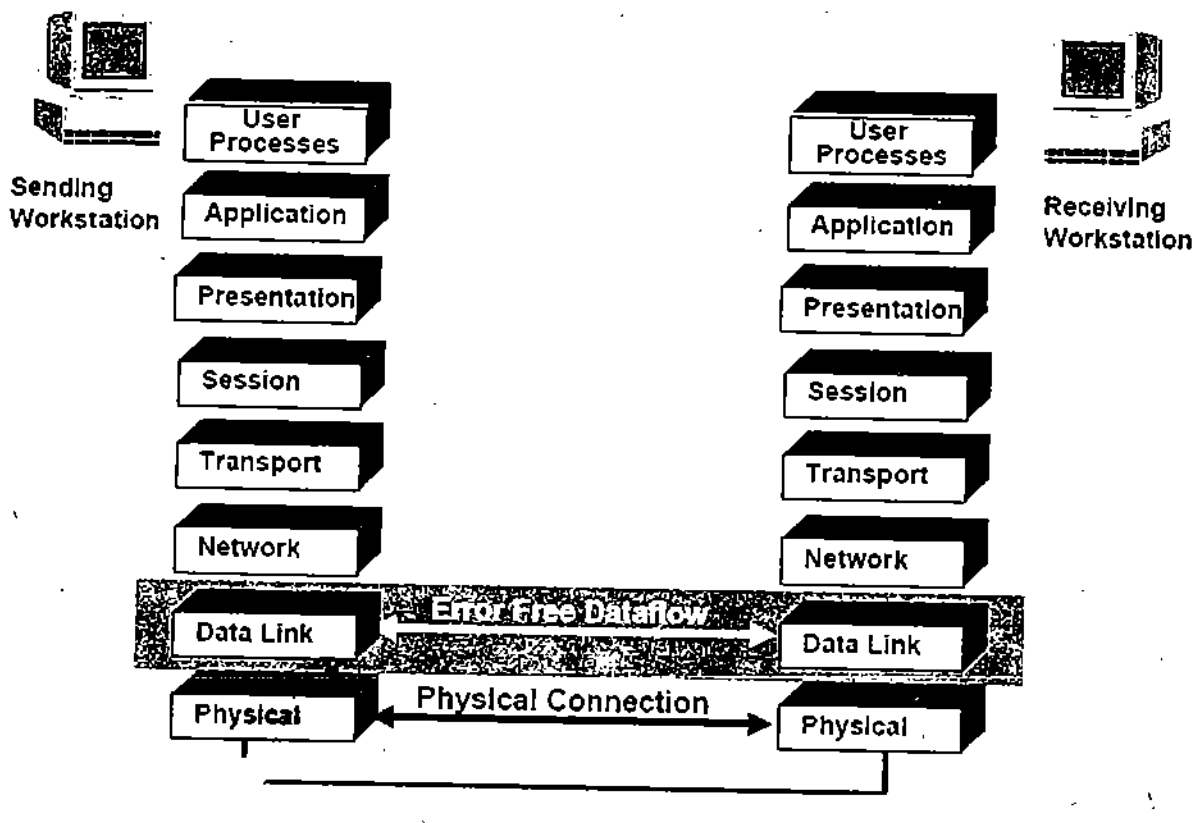
The physical layer defines:

- Physical network structures.
- Mechanical and electrical specifications for using the transmission medium (conductive cable, optical fiber or air).
- Bit transmission encoding and timing rules.



Network hardware normally associated with the OSI physical layer includes: Ethernet, twisted pair and coaxial cable, BNC connectors, RJ-45 connectors, network interface cards, repeaters, concentrators, and hubs.

The physical layer receives a frame from the data link layer above and converts the digital data to a signal compatible with the transmission media. The most common protocols between the physical layer and the data link layer are Ethernet and Token Ring.



The basic purposes of data link layer protocols are:

- To organize the physical layer's bits (0's and 1's) into logical groups of information called frames (groups of contiguous bits forming a unit of data).
- To detect (and sometimes correct) errors.
- Control data flow.
- Identify resources on the network through physical device addressing. The functions of the layers are normally split between two sub layers.
- Media Access Control (MAC) layer which controls the way devices share the same media channel. It determines when a device can transmit its data.

- **Logical Link Control (LLC)** layer is in-charge of establishing and maintaining communication between devices.

Network hardware associated with the OSI data link layer includes bridges, switches, and network interface cards (NICs).

Some protocols and standards implemented at this layer are Ethernet, PPP, FDDI, Frame Relay, and ATM.

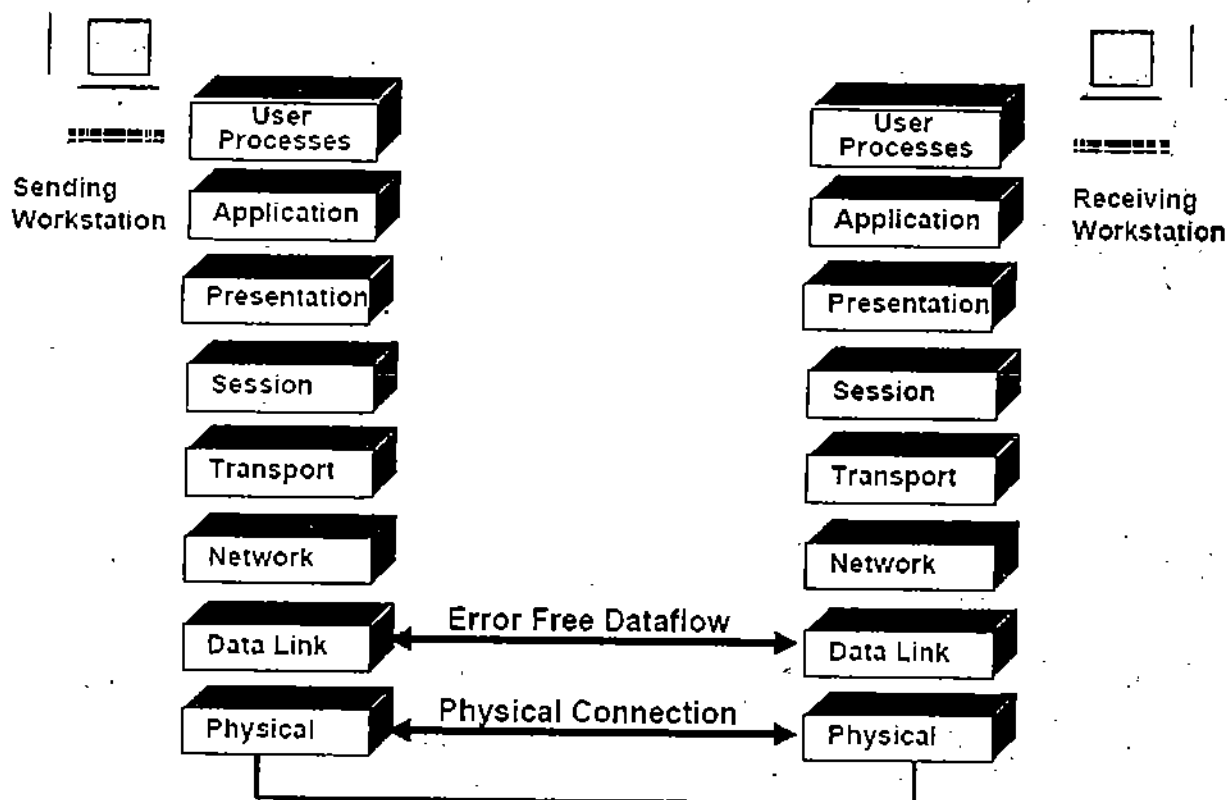
Layer 3: Network

The network layer provides the functions not found in the first two layers that allow for the exchange of information between two devices on different wires, that is, on different networks.

The Network Layer

The objectives of the network layer are:

- To move data to specific network locations. This function is similar to the

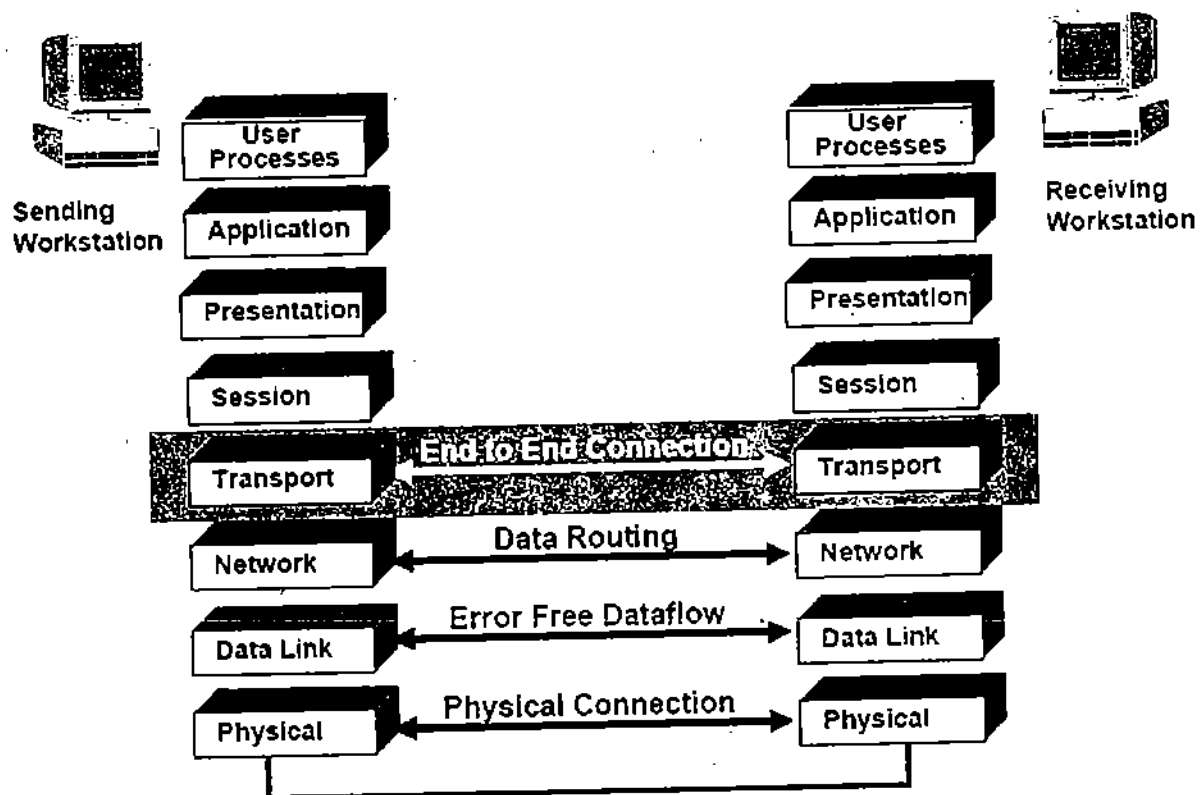


- Data link layer; however, the network layer deals with multiple independent networks, called subnetworks. For example, this layer provides the network addressing and routing structure of the Internet.
- To segment messages and deliver packets or datagrams by line switching or packet switching.
- To determine the "best" route for the data to travel from LAN to LAN.

Layer 4: Transport

The transport layer provides the functions needed to make a connection to a remote device. These transport layer protocols are designed to hide the complexities of the computer network structure from the upper layer processes. They organize higher level messages into segments and reliably deliver segments to higher layers.

The Transport Layer



The transport layer:

- Provides reliable connection-oriented (such as TCP) and connectionless (such as UDP) transmissions between end nodes.
 - Corrects transmission errors.
 - Confirms whether data was delivered or denied.
- A well-known protocol operating at this level is TCP, the Transmission Control Protocol, which is used on the Internet and is part of the TCP/IP protocol stack.

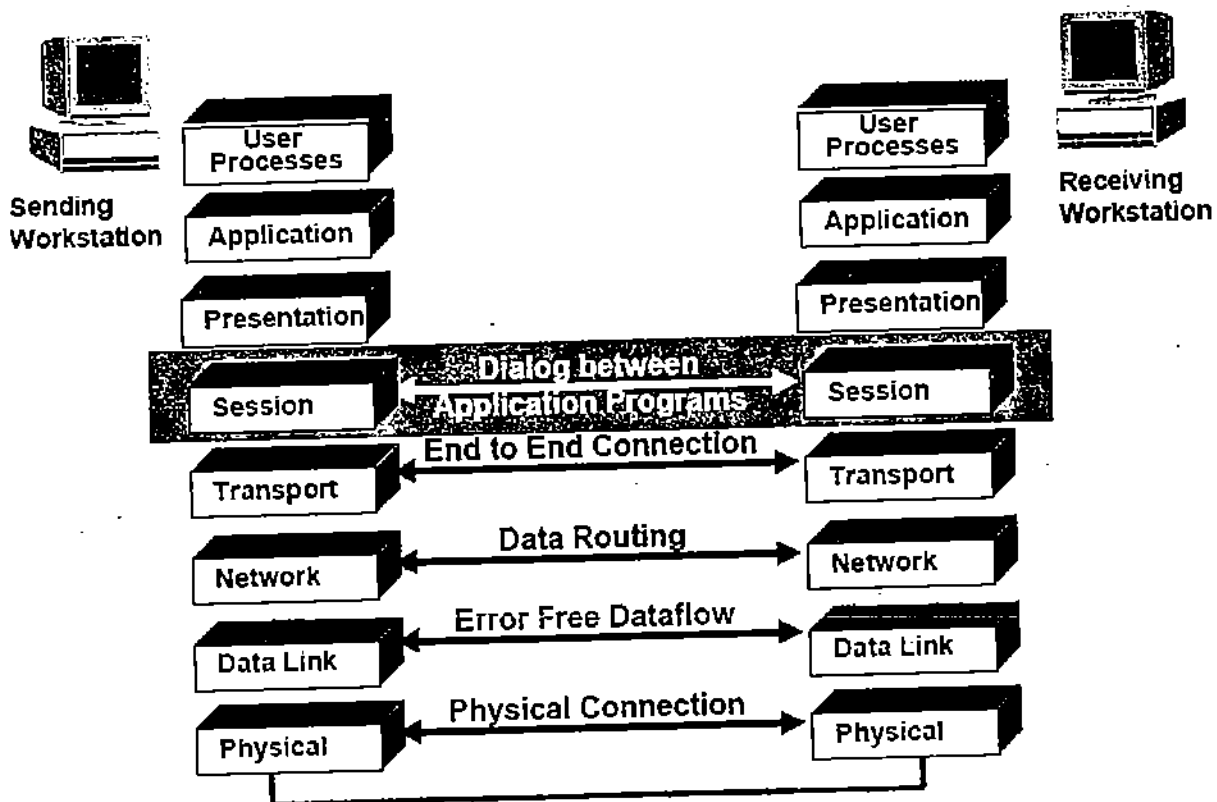
Layer 5: Session

The session layer facilitates communication between sending and receiving computers.

The Session Layer

The session layer protocols:

- Control mechanisms that establish connections and synchronize how two devices communicate.



- Provide the user interface to the network (passwords and login).
- Monitor and maintain connections between two devices, reporting error messages to users and providing monitoring and billing procedures for shared systems.

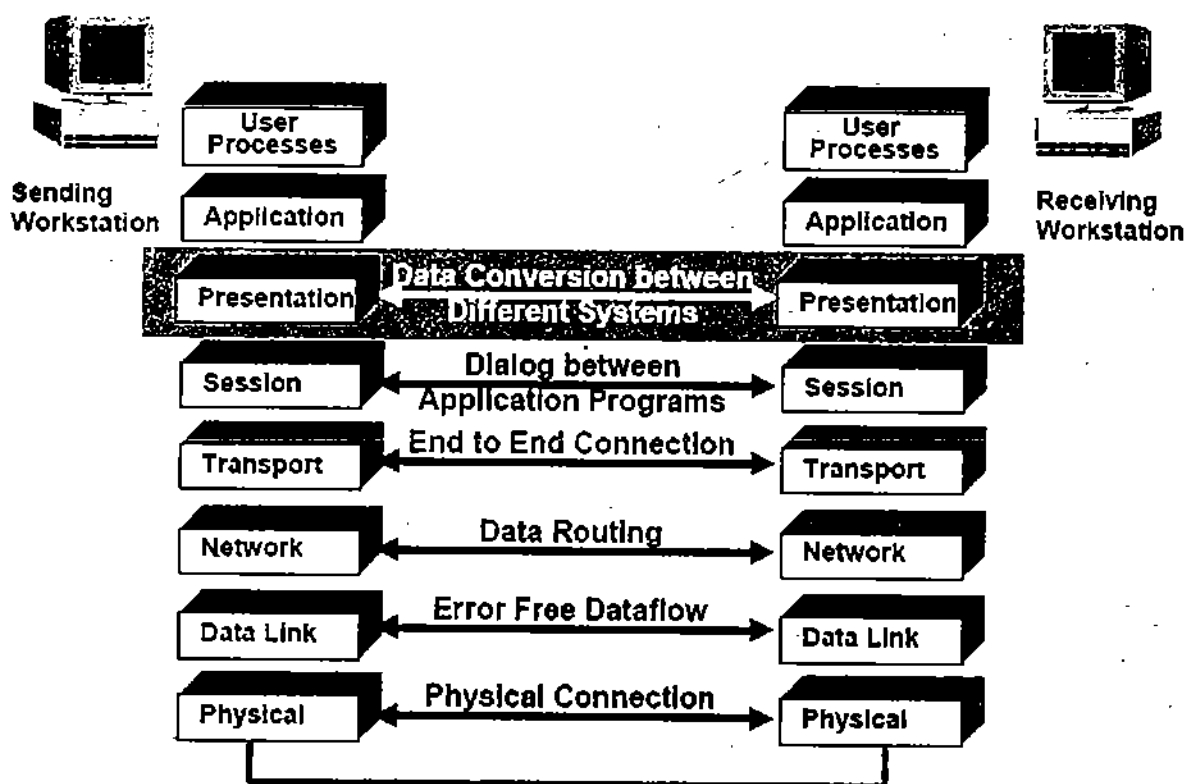
A popular personal computer session layer protocol is NetBIOS.

FTP and Telnet protocols (from the TCP/IP stack) use the connection-oriented services of this layer.

Layer 6: Presentation

The presentation layer translates data into a standard format that can be understood by each network application and the computers on which they run. The protocols in this layer also provide for compression/expansion and/or encryption/decryption of data and security services. This layer also takes care of file locking, which prevents two users from modifying the same file at the same time.

The Presentation Layer



Protocols at this level deal with rules for representation of complex, structured data across a session (for example, MIME).

Layer 7: Application

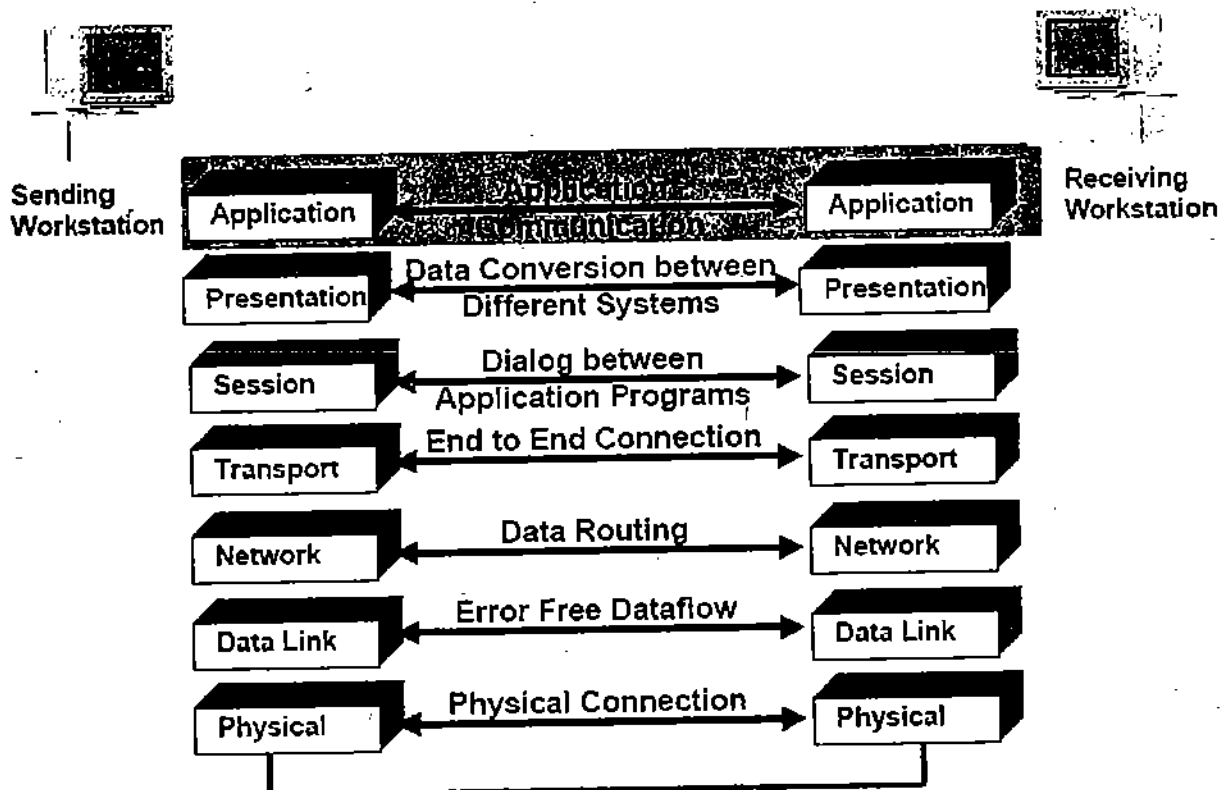
The application layer handles the management of the OSI network and provides the representation of a connection to an end user such as HTTP. The protocols in this layer also perform network functions such as file and print.

The Application Layer

This layer does not include the actual application software, but does include the technologies that allow user software to access network services such as:

- Printing and viewing.
- Interfacing databases, email, web browsers, and word processing to network resources.

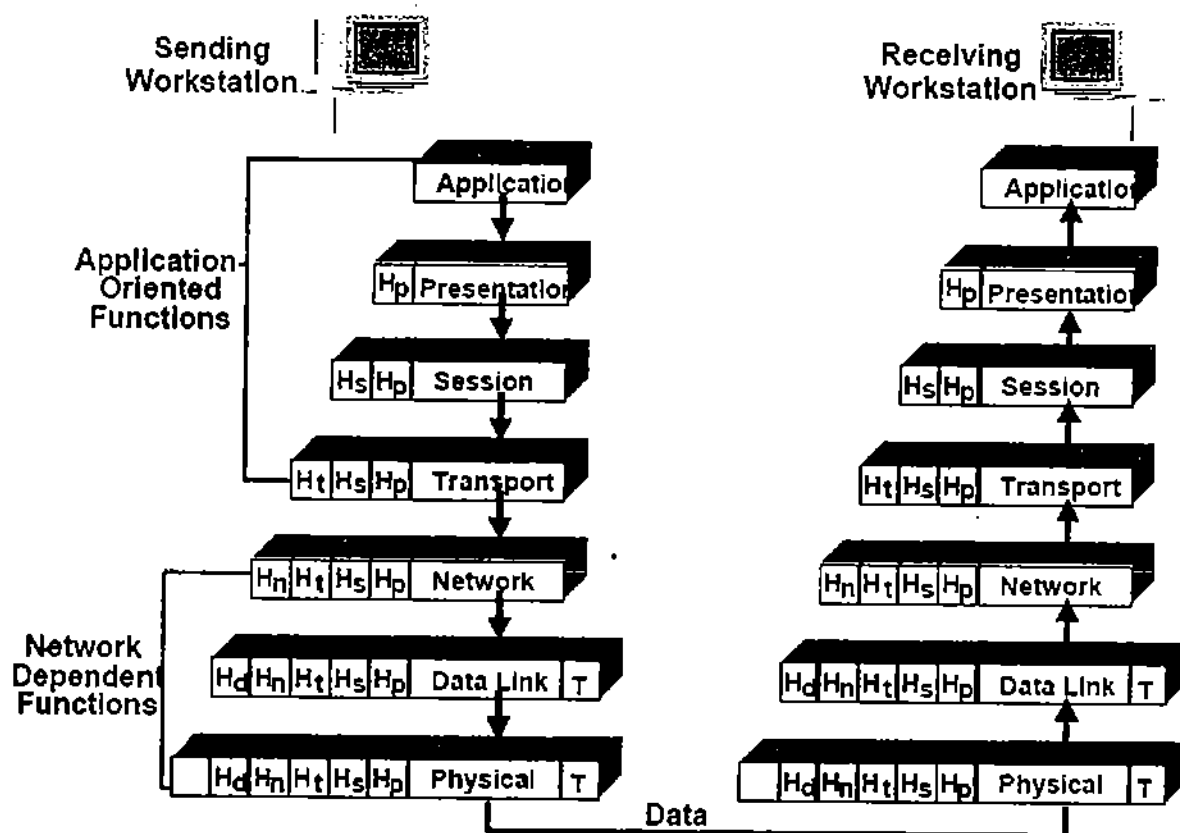
Examples of TCP/IP protocols running at this layer are HTTP and SMTP.



1.3 Data Flows Both Ways through Stacked Layers

When data is transferred from one host to another over a network, it begins its flow from the uppermost (application) layer of the originating host and continues through each lower stacked layer of that host. Each layer adds information such as framing and addressing to the data flow. This data is called a *header*. Layer 2 also adds a *trailer*. Layer 2 also adds a *trailer*.

Data Flows Through the Layers of the OSI Model



As data crosses the physical medium (Layer 1) to the destination host, the data flows back up through the layers in the opposite order, from lowest (physical) to the highest (application) layer. Each layer in the receiving system communicates with its peer layer in the sending system by reading and processing the (header) control information. When successive layers remove all headers, the receiving system sees the data in original format.

It is interesting to note that as data flows from one host to another, each layer functions as if it alone is involved in communication with its peer layer at the destination host. Individual layers are unaware of the functions of layers above and below them.

OSI Layers Work as a Stack

The seven OSI layers work together as a stack to encode and send messages through an open system. Peer layer protocols in another connected network or network device then process these messages.

1.4 Summary

This unit teaches the importance of OSI Model. It describes seven layers of the OSI model and how they communicate with one another. It also describes the functions performed by each of the OSI Model layers. Then we learn how to match network elements with their appropriate OSI layer. Finally, an understanding of the OSI Model is made in context of information technology carriers.

1.5 Review Questions

OSI Model

Part A

1. Explain why the OSI model is important.

Part B

1. List the seven layers of the OSI model in stacked order (Layer 7 at the top)
2. Describe how the layers of the OSI model communicate with each other.

Part C

1. Match the OSI layer with the appropriate function from the list below.

Layer				Function
7 – Application			A.	Routes and addresses messages
6 – Presentation			B.	Translates data so software applications can read it
5 – Session			C.	Allows user software to access network services
4 – Transport			D.	Enables network logins
3 – Network			E.	Frames and controls data flow
2 – Data Link			F.	Provides reliable transmission and confirmation
1 – Physical			G.	Defines media specifications

Part D

1. Label network hardware elements and functions with the appropriate OSI layer.

Layer	Device and Function
	Repeater~a connector that filters, amplifies and retransmits information to allow signals to travel further along a network.
	Bridge~connects small numbers of similar LANs into an internetwork or splits an overloaded network into smaller parts without translation.
	Router~connects different networks together and directs information by best route (using algorithms) to destinations.
	NIC~provides connection of computer to transmission medium on the network.
	Gateway~links different network types together. Can be hardware or software that allows different protocols to exchange information.
	Hub~provides central location with ports where cables on a network come together for connection
	Modem~allows computers on a network to exchange information by translating into binary form for transmission over telephone lines.
	Cable~wires used to connect devices on a network.

Unit III - Ethernet

What You Will Learn (Objectives)

After completing this unit you will be able to:

Describe Ethernet topology.

Explain how CSMA/CD is used within an Ethernet LAN.

Compare and contrast two Ethernet frame types.

Describe the relationship between IEEE Ethernet standards and the OSI model.

Troubleshoot an Ethernet problem.

Terminology

- **Back Off Mode**~ A waiting mode in Ethernet topology where computer devices wait a specific amount of time before attempting to retransmit data after they have detected a collision.
- **Bandwidth**~ The amount of data a transmission medium can carry.
- **Baseband**~ A data transmission method in which the entire bandwidth of the cable is used to transmit a single signal. Only one signal can be transmitted at a time.
- **Broadband**~ A data transmission method in which cable capacity is divided into multiple independent bandwidth channels. This allows several data transmissions to occur simultaneously over the cable.
- **Bus**~ A bus is a common pathway (usually copper wire or fiber-optic cable) between multiple devices such as computers. A bus is often used as a backbone between devices. **Sense**~ The ability to detect or sense the presence of other data transmissions on a network. In Ethernet technology, this ability helps avoid or prevent collisions.
- **Collision Detection**~ The ability to detect whether any other devices are trying to access the network simultaneously.
- **Collision Domain**~ All devices on the same Carrier Sense Multiple Access/Collision Detection wire segment are considered one collision domain. A collision in one domain does not affect any other domain on the network.
- **Contention Based**~ In Ethernet topology, a computer device must wait for a turn to use the network. Computer devices must contend with other devices for network access. This means that all devices have an equal opportunity to send signals. No device has priority.
- **Ethernet**~ A LAN topology based on Carrier Sense Multiple Access/Collision Detection (CSMA/CD).
- **Network Operating System (NOS)**~ Software that allows computer devices to

interconnect in order to transmit data across media.

- **Passive Hub**~ A central connecting device in a network that joins wires from several stations, but does not provide signal processing or regeneration.

1.1 Ethernet Origins

Ethernet topology, which is based on bus and bus-star physical configurations, is currently the most frequently configured LAN network architecture. A bus is a common pathway (usually copper wire or fiber cable) between multiple devices such as computers. A bus is often used as a backbone between devices. It is a technology that has been evolving for more than 25 years and is still evolving to meet the ever increasing and changing needs of the internetworking community.

Digital Equipment Corporation and Xerox (DIX) worked together to develop the first Ethernet standards. These standards are the DIX Ethernet standards and are still in use today. As Ethernet topology became more popular, industry-wide standards became necessary. In 1985, the IEEE adopted the current Ethernet standards. These standards are called the IEEE 802.2 and 802.3 standards. They differ slightly from the DIX standards, but both define the protocols for the physical and data link layers of the OSI Model. These standards include cabling specifications, frame format, and network access conventions.

Ethernet is a passive, contention-based broadcast technology that uses baseband signaling. Baseband signaling uses the entire bandwidth of a cable for a single transmission. Only one signal can be transmitted at a time and every device on the shared network hears broadcast transmissions. Passive technology means that there is no one device controlling the network. Contention-based means that every device must compete with every other device for access to the shared network. In other words, devices take turns. They can transmit only when no other device is transmitting.

Ethernet popularity is a result of several factors. Ethernet technology is:

Inexpensive

Easy to install, maintain, troubleshoot and expand

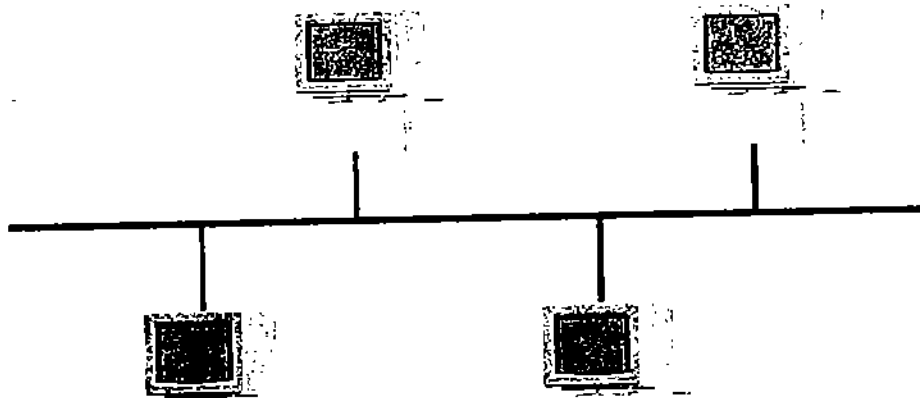
A widely accepted industry standard, which means compatibility and equipment access are less of an issue

Structured to allow compatibility with network operating systems (NOS) Very reliable

1.2 Ethernet Configuration and Communication Ethernet Configuration

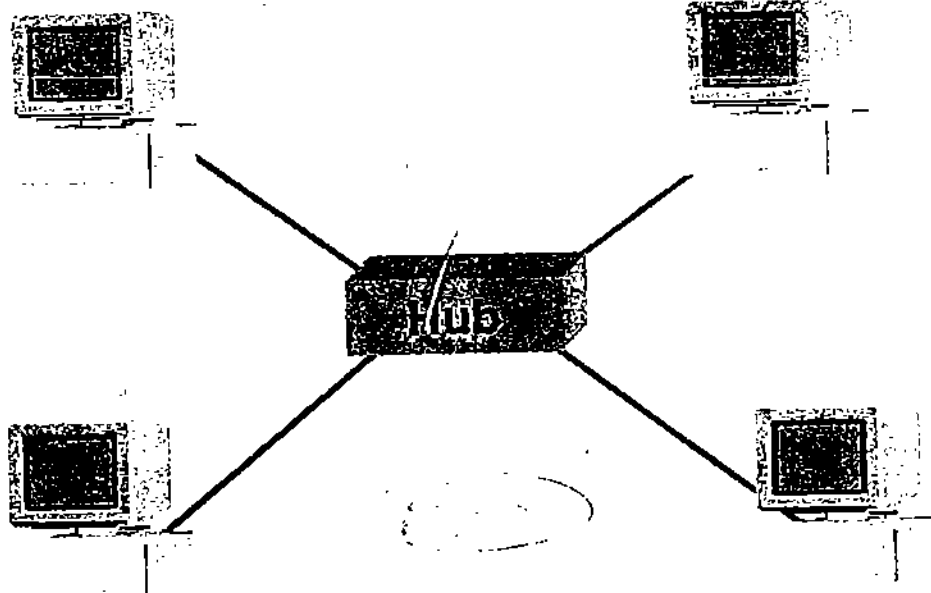
Ethernet is a broadcast topology that may be structured as a physical bus or physical star with a logical bus.

Ethernet Physical Bus Topology



The Physical star with a logical Bus is created with the use of a hub or concentrator.

Ethernet Physical Star/Logical Bus Topology



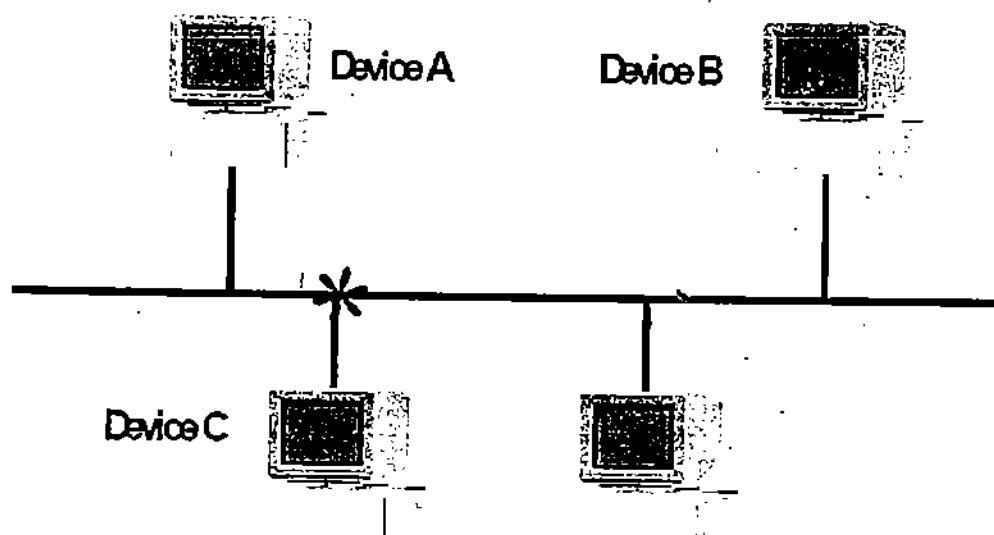
1.3 Ethernet Communication

Communication protocols for Ethernet networks encompass both the data-link and physical layers of the OSI model. This lesson deals mainly with the data-link layer, which is subdivided into a Media Access Control layer and a Logical Link Control layer. Lesson 2-4 covers physical layer issues.

Ethernet uses Carrier Sense Multiple Access/Collision Detection (CSMA/CD) when transmitting data. Carrier Sense allows a computer device to "sense" whether or not another transmission is being "carried" over the network. So, before a device sends data, it listens for a carrier (jam) signal. If a carrier signal is detected, it waits until that transmission is completed. Early DIX Ethernet did not have a carrier signal. Therefore, a collision was not detected until the destination device received the framed packet. The addition of a jam signal is one example of how Ethernet technology has evolved.

Multiple Access means that all devices have equal access to the network. Since Ethernet is contention-based, equal access to the network for all is ensured. No device has priority over others, nor can it lock out any other device connected to the network. Information can be transmitted at any time by any device. All devices on the network receive the transmission and check the framed packets destination address. If the destination address matches the device's address, the device accepts the data; if the address does not match, the device simply ignores the transmission.

1.4 Ethernet Collision



Collision Detection means that a sending device can "detect" simultaneous transmission attempts. When two or more devices try to send data at the same time, the signals collide. The illustration above shows devices A, B, and C sending

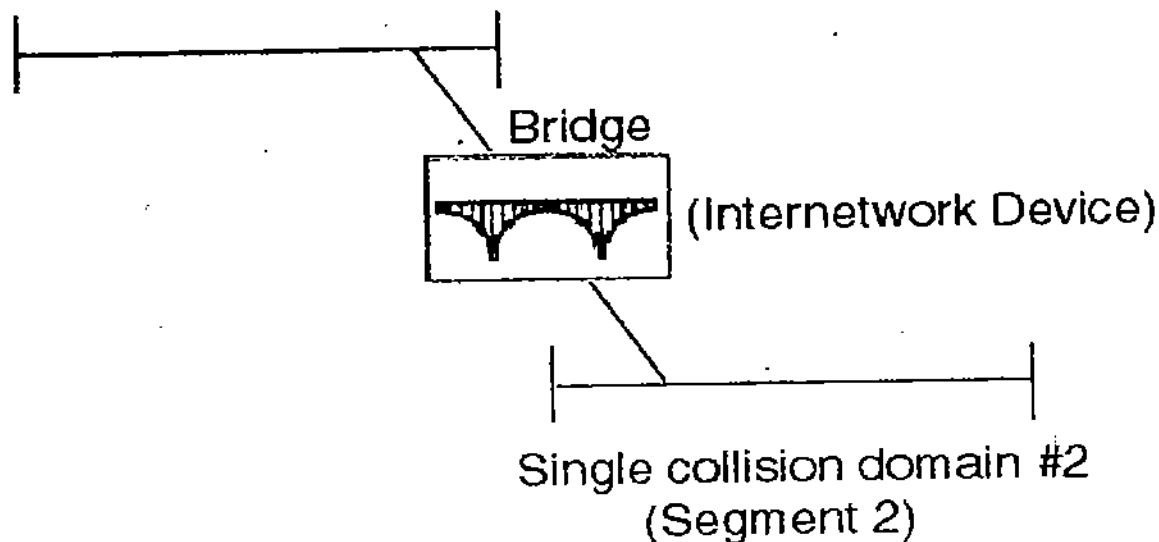
signals simultaneously, and a collision occurs. When this happens, each device then transmits a jam signal, called a carrier, to alert all other devices that a collision has occurred.

All devices then go into back off mode and wait a random amount of time before attempting to retransmit data. The random time provision prevents simultaneous retransmissions. All devices on the same Carrier Sense Multiple Access/Collision Detection wire segment are part of the same collision domain. A collision domain is defined as those devices that share CSMA/CD of the same wire. Two or more collision domains are connected together with an internetworking device such as a router, bridge, or switch. With the use of internetworking devices, large networks are created which include multiple collision domains.

When a collision occurs it affects all the devices on the same collision domain. It does not affect devices on any other collision domains within the same network. You can think of two collision domains as two roads that are on different sides of a bridge. You can travel up to the bridge on either side, but must get permission to cross the bridge. On an Ethernet collision domain, frames of data travel within **their** own domain unless they need to talk to a device on the other side of the bridge. If it is necessary to talk to a device on the other side of a bridge, the bridge must give permission for the frame to cross the bridge to reach the other collision domain. Bridges are capable of this because they are store and forward devices: they store the frame from the source domain until permission is granted to forward it. The requirement is to keep traffic and collisions to a single collision domain whenever possible.

Ethernet Collision Domain

Single collision domain #1
(Segment 1)



Data transmitted from one device to another on the same collision domain will not affect any other collision domains. This allows each collision domain to continue to transmit with no effect on each other except when a device needs to talk to a device on another collision domain. When this is required, the frame must be sent across one or more internetworking devices to reach its destination. Each internetworking device must allow the frame to pass. As the frame reaches each collision domain the port of the internetworking device must contend for the right to transmit according to the rules of CSMA/CD.

To better understand CSMA/CD, think about trying to make a telephone call. Many of us have more than one telephone in our homes (a telephone network). When you pick up the telephone to make a call, you "sense" a dial tone or someone else on the line. If there is a dial tone, you proceed with your call. If the telephone line is currently in use, you can not make a call at this time and you try again later. This is similar to Ethernet Carrier Sense protocols.

All telephones in the house can be used at any time to make calls. All phones in the house have equal access to the telephone network. This is comparable to Multiple Access.

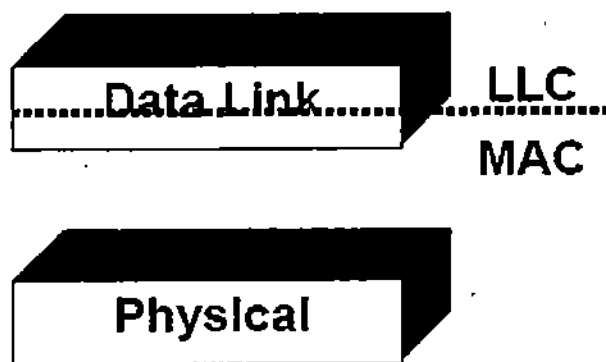
Should two individuals in the house attempt to make a phone call simultaneously, both hear dial tone; neither party senses a carrier, (someone else on the line). However, like Ethernet technology, only one individual can use the line at any one time. Both parties must hang-up and wait a random amount of time before making a second attempt. This is how Ethernet's Collision Detection protocols work.

In summary, a computer device checks to see if the transmission media is busy, recognizes that multiple devices access the network, and detects when a collision has taken place and goes into back off mode.

1.5 Ethernet Frames

In Ethernet, both the data link and the physical layers are involved in the creation and transmission of frames. The physical layer is related to the type of LAN cabling and how the bits are transmitted and received on the cable. The data link layer is divided into sublayers, the Logical Link Control (LLC) and the Media Access Control layers (MAC). The frames created by these layers contain several fields that are processed by Network Interface Cards (NICs) in the sending and receiving devices.

Ethernet and OSI Layers



The MAC sublayer address is the physical hardware address of the source and destination computer. It is called the MAC layer address and should not be confused with the network address. All devices on a LAN must be identified by a unique MAC address. This sublayer controls which computer devices send and receive the data and allows NICs to communicate with the physical layer. IEEE 802.3 protocols control the format of the MAC sublayer frame fields.

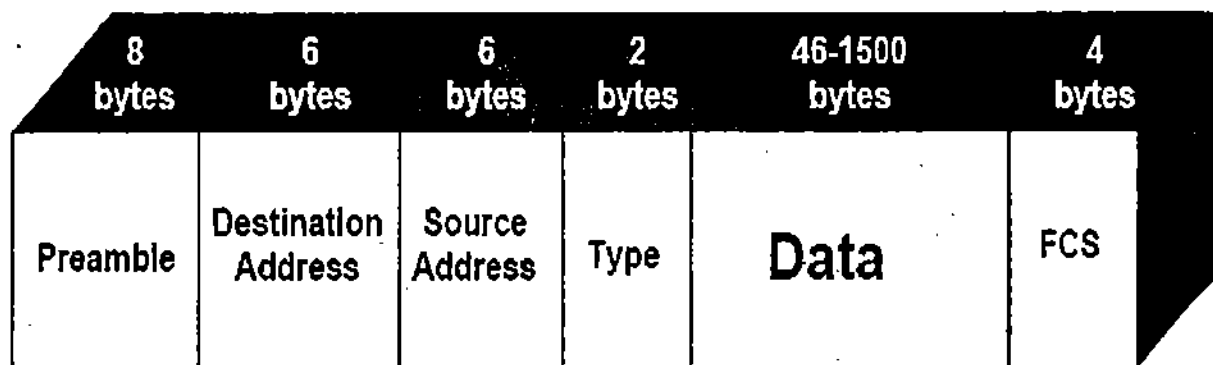
The next level of processing is the LLC sublayer. It is responsible for identifying and passing data to the network layer protocol. Two LLC protocols are IP and Novell's IPX.

1.5.1 Frame Format

The two frame formats discussed in this lesson are DIX frames, which are what the Internet uses, and IEEE 802.3 frames. It should be noted that if one device uses an 802.3 NIC and the other device uses a DIX Ethernet NIC, they would not be able to communicate with one another. Devices must create the same Ethernet frame format in order to be compatible. Although only the 802.3 IEEE frames format is outlined in the standard, both formats are in use today. One way to tell them apart is that the DIX frame has a "type" field, which defines the protocol used for the frame, and IEEE 802.3 has a "length" field in its place. IEEE 802.3 also has additional fields not used with the DIX format.

The DIX, also called Ethernet II, frame includes the following fields:

DIX Frame (Ethernet II)



The Preamble of the frame (the first 7 bytes) indicates the start of a new frame and establishes synchronization conditions between devices. The last byte, or start frame delimiter, always has a 10101011-bit pattern. This byte indicates the start of a frame.

The Destination Address is the hardware (MAC) address of the receiving device, and the source address specifies the hardware (MAC) address of the sending device.

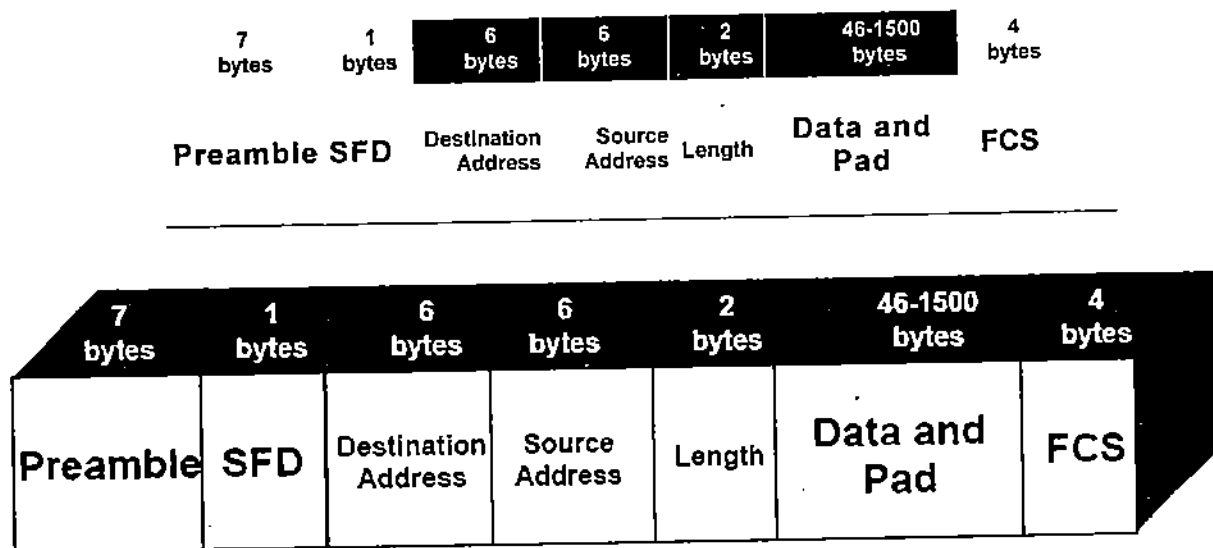
The Type field specifies the network layer protocol used to send the frame, for example TCP/IP.

The data field is for the actual data from device to device. It also contains

information used by the network layer and indicates the type of connection.

1.5.2 IEEE 802.3 Frame Format

Fields one and two perform the same function as the DIX preamble; however, the fields are separate. The Start Frame Delimiter (SFD) has the same 10101011-bit sequence found at the end of the DIX preamble. Both formats use the same number of bytes to perform the synchronization of the signals.



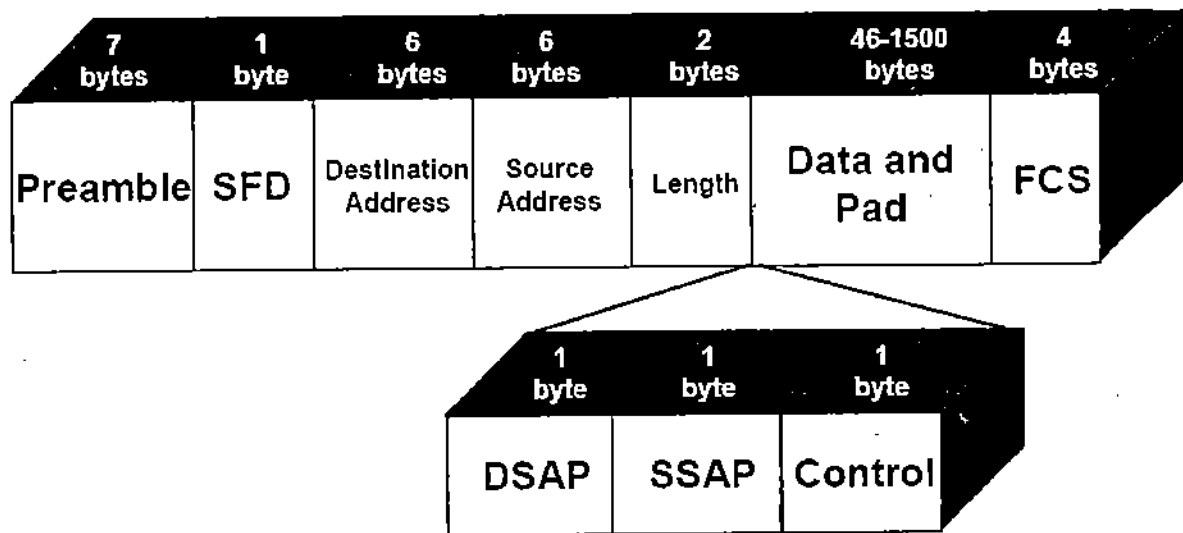
The Destination and Source Addresses can be either 2 or 6 bytes. Whether 2 or 6 bytes are used, all devices within the same network must use the same format. IEEE protocols specify that a 10Mbps network must use 6 bytes. The 2 byte length is obsolete.

The Length field indicates the number of bytes in the data field. If the data field is less than the required 46 bytes, a pad field is added to the data frame. The bytes added for padding purposes are usually zeros.

The **data** field contains the data to be transmitted from device to device.

The Frame Check Sequence (FCS) field is used as an error detection function. The error detection function is a calculation completed by both the source and destination devices. If the calculations do not match, an error is then generated.

1.5.3 IEEE 802.2 LLC Frame

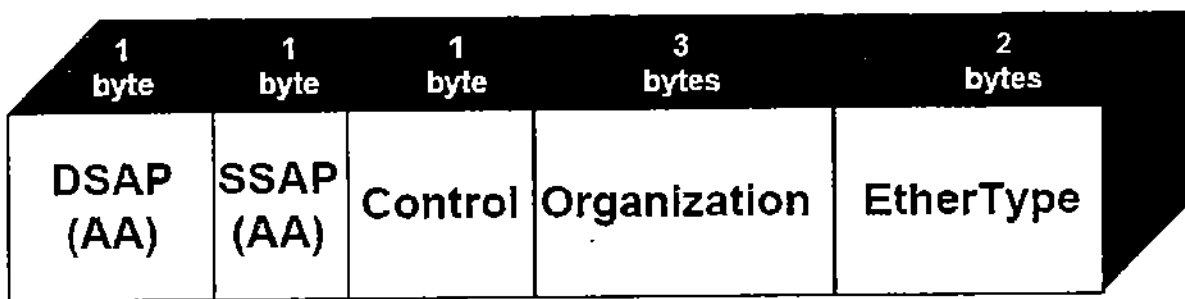


In between the length field and data/pad field, is the 802.2 LLC field.

DSAP SSAP Control

- DSAP, destination service access protocol, is the protocol processing layer the data is to be sent to.
- SSAP, source service access protocol, is the protocol used to encapsulate the data the source.
- Control is the field that defines the type of field this is.

SNAP Format of 802.2



SNAP (Sub-Network Access Protocol) was created by the IEEE to identify the Network layer protocol used. The original Ethernet version 2.0 Type field had been reused as a Length field by the IEEE when creating the IEEE 802.3 standard. SNAP was defined to enable this Length field to remain while also allowing the vendor and protocol to be defined in the first 40 bits of the Data field. Together, these two fields (Organization and Ether Type) are called the Protocol ID.

AA in the DSAP or SSAP fields indicates that the LLC field is using SNAP format.

Control is the field that defines the type of LLC frame this is.

Organization is the field that indicates which organization created the protocol identified in the Ether Type field, though generally this is coded as all zeros by most organizations.

Ether Type is a two-byte identifier for the protocol being used to encapsulate the data. For example, IP is indicated by the code 0x08-00, ARP by 0x08-06 and NetWare IPX/SPX by 0x81-37.

Depending on how the vendor decided to manufacture the Ethernet hardware, the hardware will create frames that are either DIX, 802.3, or 802.2 with SNAP.

Frame Types

There are three types of frames; each has a different purpose. The three types are:

- Unicast
- Multicast
- Broadcast

If the first bit of the frame is 0, it is Unicast; if it is 1, it is multicast. Broadcast frames always have 1 as the second bit.

A Unicast frame is addressed to a single network device. This means that the frame is to be read only by the device that matches the destination address. All other devices in the collision domain will receive a copy of the frame but will discard it

because it does not match their destination address. The address used is the MAC address or hardware address of the network device.

A multicast frame is addressed to several but not all devices. All devices that are a part of the specified group may read the frame. A multicast address is a deviation from the normal hardware address. For example, a group of devices are assigned access to a particular server on the network. They are the only devices that receive frames announcing the availability of that server. Any device that does not belong to this group will ignore or discard these frames.

A broadcast frame is addressed for all network devices to read and process. A broadcast address is a unique address used only for broadcast frames. It is not a hardware address. Broadcast frames are transmitted across bridges and switches; however, routers will stop broadcast frames.

1.6 Summary

In this unit we learned to do the following:

- Describe Ethernet topology.
- Explain how CSMA/CD is used within an Ethernet topology.
- Compare and contrast two Ethernet frame types.
- Describe the relationship between Ethernet standards and the OSI model.
- Set up an Ethernet LAN and ping to check for connectivity.
- Troubleshoot an Ethernet problem scenario.

1.7 Review Questions

Part A

1. Which type of transmission do all devices on the shared network hear and accept?
 - a) Unicast
 - b) Multicast
 - c) Broadcast
 - d) Baseband

2. Which type of signaling uses the entire bandwidth of a cable for a single transmission and allows only one signal at a time?
 - a) Unicast
 - b) Multicast
 - c) Broadcast
 - d) Baseband

3. When all devices have equal access to the network and no one device has priority over another device, what is this called?
 - a) Carrier Sense
 - b) Multiple Access
 - c) Collision Detection
 - d) Collision Domain

4. What is the term used to describe the ability of a device to sense simultaneous transmission attempts and wait a random amount of time before retransmitting data?
 - a) Carrier Sense
 - b) Multiple Access
 - c) Collision Detection
 - d) Collision Domain

5. What is the ability to listen for a jam signal before transmitting data called?
 - a) Carrier Sense
 - b) Multiple Access
 - c) Collision Detection
 - d) Collision Domain

6. Ethernet standards include specifications for which of the following?
 - a) Cabling
 - b) Frame format
 - c) Network access conventions

d) All of the above

7. What are the two IEEE Ethernet standards called?
8. What is the name for the Ethernet standards developed by Xerox, Intel, and Digital Equipment Corporation?
9. What type of topology configuration(s) does Ethernet employ?
10. List five reasons for the popularity of Ethernet LAN topology.

Matching: Match the part of the frame with its definition.

1. ___ Preamble	A	One byte code in the LLC field used to identify the protocol that will encapsulate the data field
2. ___ Start of Frame	B	Hardware address of the destination device
3. ___ Destination Address	C	Actual information being transmitted
4. ___ Source Address	D	Specifies the protocol used for sending the frame
5. ___ Type Field	E	Specifies the length of the data within the frame
6. ___ Length Field	F	Type of frame check that detects errors that occur in the frame during transmission
7. ___ Pad	G	Added to the data field of IEEE 802.3 when the data is less than 46 bytes
8. ___ Data	H	Establishes synchronization and transceiver conditions
9. ___ DSAP	I	Field with 10101011 sequence separate
10. ___ CRC	J	Hardware address of the sending Device

Unit IV – Networks & Devices

What You Will Learn (Objectives)

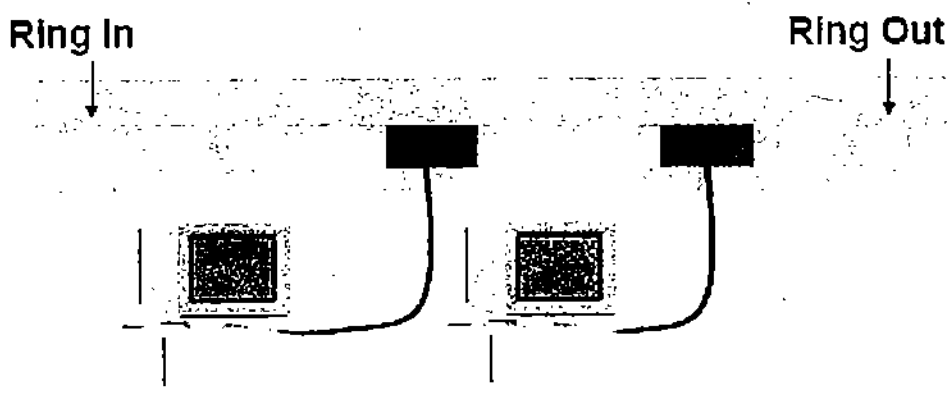
After completing this unit you will be able to:

- Describe the characteristics of Token Ring architecture.
- Describe the characteristics of FDDI architecture.
- Design a Token Ring network
- Describe the characteristics of ATM
- Identify network connectivity and internetworking devices.
- Match the appropriate level within the OSI model to which each device is assigned.
- Compare the functionality of connectivity and internetworking devices

4.1 Token Ring Architecture

Token Ring architecture, introduced in the mid 80s by IBM and defined by the IEEE 802.5 standard, is a physical star, logical ring topology. In Token Ring networks, computer devices are connected logically in a ring with each device connected separately to a specialized Token Ring hub called a multi station access unit (MAU). The function of this hub is to ensure that the packets of data are transmitted around the ring and have the ability to bypass a device in case of failure on one of the ports. Token ring hubs have two additional ports called Ring In and Ring Out that are used for connecting MAUs together.

Token Ring Multi station Access Unit (MAU)



Traditional Token Ring networks transmit signals at a speed of either 4 or 16 Mbs using a token that is passed from computer to computer around a network. [Recent Token Ring technology with speeds of 100 and 128 Mbs is called fast Token Ring networks. These networks operate like traditional Token Ring networks]. In order to send data, a computer device must capture the token. The token, which is continuously transmitted around the ring, is a 24-bit frame with three fields, a Start Delimiter (SD), an Access Control (AC), and an Ending Delimiter (ED).

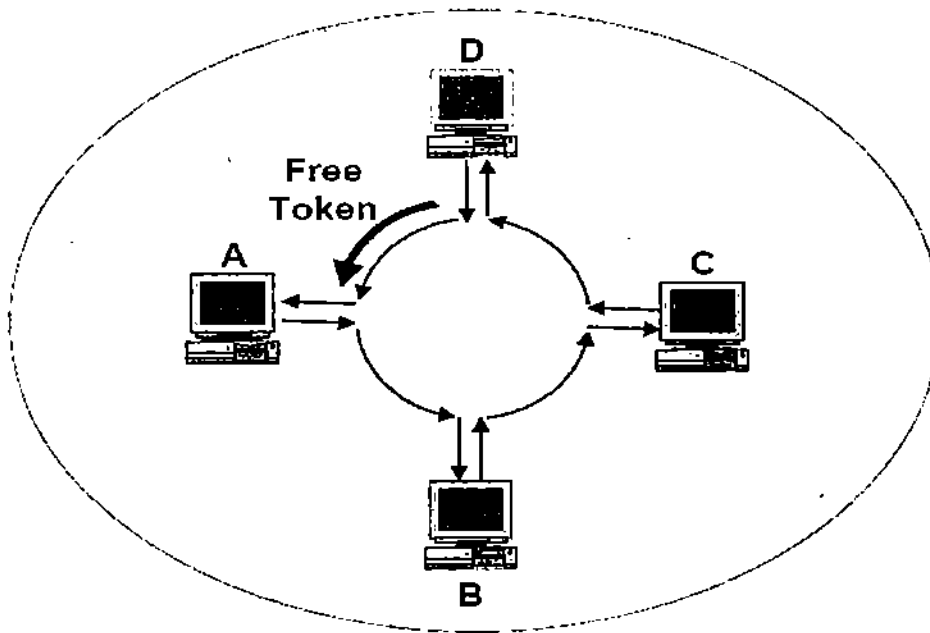
Token Frame Format

SD	AC	ED
----	----	----

The access field carries a signal, either a 0 or 1, that tells the computer device whether or not it is available to carry data. If available, the access control signal is a 0, and the computer device captures the token frame, inserts the information being transmitted in between the access field and the ending delimiter, and then sends the entire frame.

When workstation A transmits to workstation C, workstation A grabs a free token and changes the AC to a 1 indicating that the token is no longer available for use and encapsulates its data within the token frame

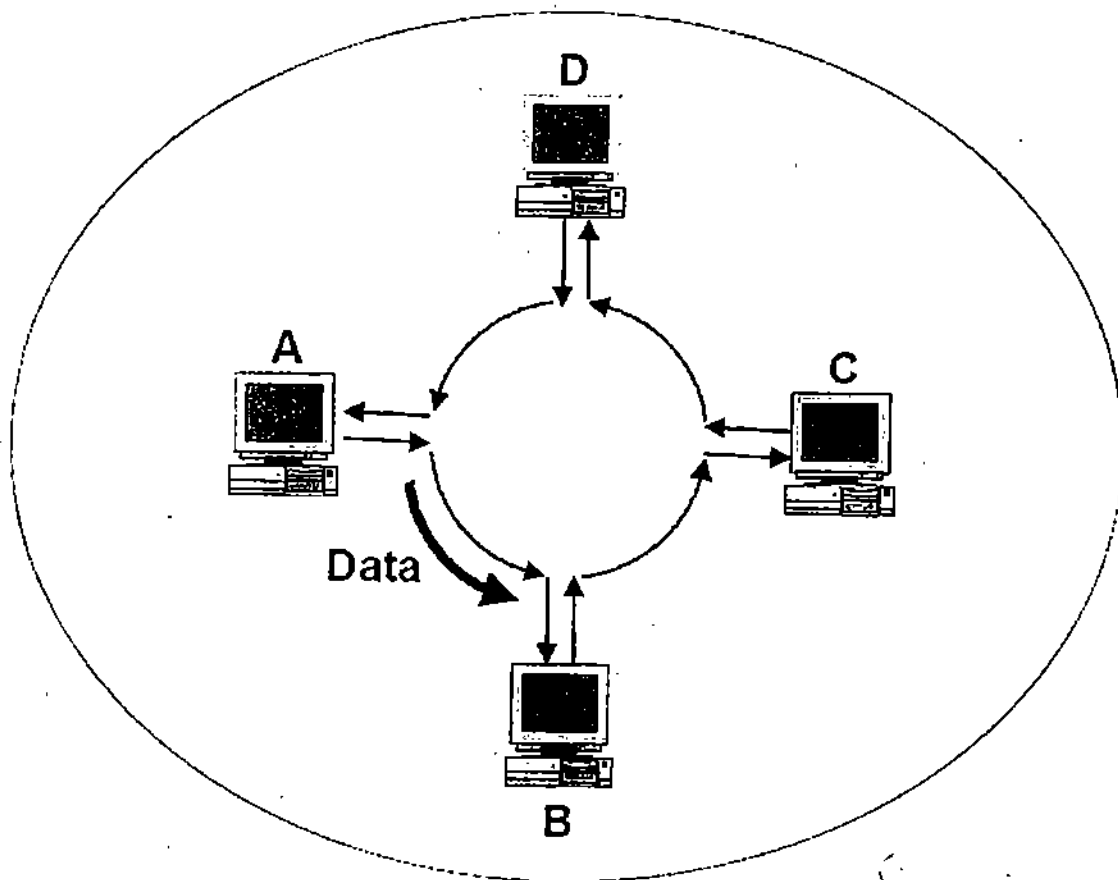
Workstation A Captures the Free Token



The token with its encapsulated data travels around the network from one computer to the next, where each device accepts the frame and checks the destination address. If the hardware addresses match, the device processes the data; if it does not match the data is retransmitted to the next device until the token reaches its intended destination.

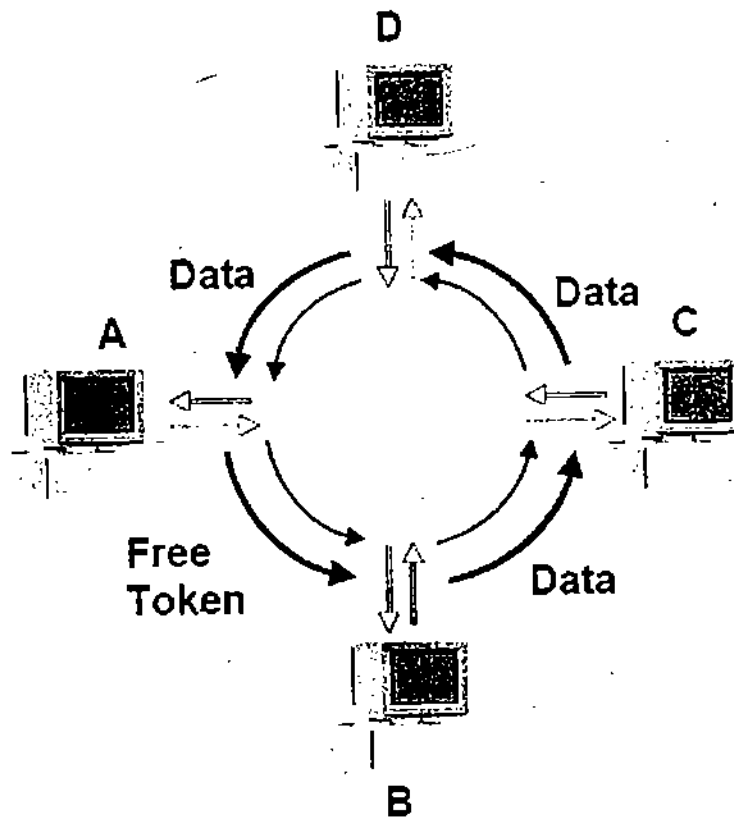
Data

Encapsulated in Token Frame



The destination device accepts the frame and copies the data, verifies receipt, and returns the token frame to the network where it continues around the ring until it reaches the original sending device. The sending device recaptures the token that contains information that the frame was received and also contains error control information. After stripping the original data from the frame, the sending device returns the available token to the network and the process continues.

Workstation C Accepts Data and Sends it Back to A to be Stripped



Four (4) Mbs Token Rings only allow one token on the ring at a time, sixteen (16) Mbs Token Rings have a mechanism called Early Token Release, which allows a free token to be released before it has been completely received by the source station.

Token Ring networks are not as susceptible to signal degradation as other network types, since each device acts like a repeater, regenerating the signal before transmitting it to the next device. Token Ring networks fail more gracefully than Ethernet networks when they are overloaded. Since they are so reliable, diagnostic and troubleshooting measures are often built into the hardware. Also, priority can be assigned in Token Ring networks.

Although they work quite well, Token Ring networks are currently being phased out for several reasons. They have very complicated protocols, which makes them both difficult to troubleshoot and more expensive than other technology. Token Ring is basically IBM, so there are compatibility issues with equipment from other manufacturers who also are not as interested in new Token Ring technology since the market isn't as profitable as the Ethernet market. Additionally, 4 Mbps networks are considered slow when compared with 10Mbps Ethernet networks.

Token Ring Frame Format

Token Ring frame format begins with the 8 bit Start Delimiter followed by the 8 bit Access Control. The next field, is a Frame Control field, which indicates whether the frame contains Media Access Control information for all network devices or only data for a single destination device. This field is followed by the Destination Address field (either 16 or 48 bits), the Source Address field (either 16 or 48 bits), the Data field (has no determined size), the Frame Check Sequence field (32 bits), the End Delimiter field (8 bits), and finally the 8 bit Frame Status field. [Note: current Ethernet LANs use the 48bit source and destination addressing scheme.]

Token Ring Frame Format

8 bits	8 bits	8 bits	16 or 48 bits	16 or 48 bits	>0 bits	32 bits	8 bits
SD	AC	Frame Control	Destination Address	Source Address	Data Field	FCS	ED

Detecting Errors through Beaconsing

One computer device on a Token Ring network, usually the first device recognized when the LAN comes up, is designated the active monitor. It is the responsibility of the active monitor (AM) to ensure that the network is functioning properly. Every seven seconds the active monitor (AM) station sends a signal/frame, which checks for malfunctioning devices and problems such as errors in delivery of frames. Each successive device on the ring receives and forwards the active monitor present frame. If a device does not receive the active monitor (AM) present frame, it sends a message/frame on to the network indicating that there is a malfunction, and the Token Ring tries to self-correct the problem. If the active monitor (AM) malfunctions, its nearest active upstream neighbor (NAUN) becomes the new active monitor (AM) and sends out a Claim Token notifying all other devices of the change in active monitor status.

4.2 Fiber Distributed Data Interface (FDDI)

Fiber Distributed Data Interface (FDDI) is a high speed (100 Mbs), fiber optic LAN technology that uses the token passing method to access the network. The American National Standards Institute (ANSI) developed standards for FDDI networks (X3T9.5) that address the Physical and Media Access Control sub-layers of the OSI model. It assumes the use of IEEE 802.2 Logical Link Control sublayer standard above the MAC sub-layer. FDDI uses physical star, logical ring topology. FDDI is a dual ring topology that is commonly used between LANs in metropolitan area networks to provide high speed connections.

Like Token Ring, FDDI uses a token frame to pass data from one station to the next around a ring until it returns to the source device; however, an additional clocking field is added to the beginning of the token. This field is added because FDDI transmissions are synchronous and require a clocking mechanism. Voice and video transmissions are examples of data that require synchronization.

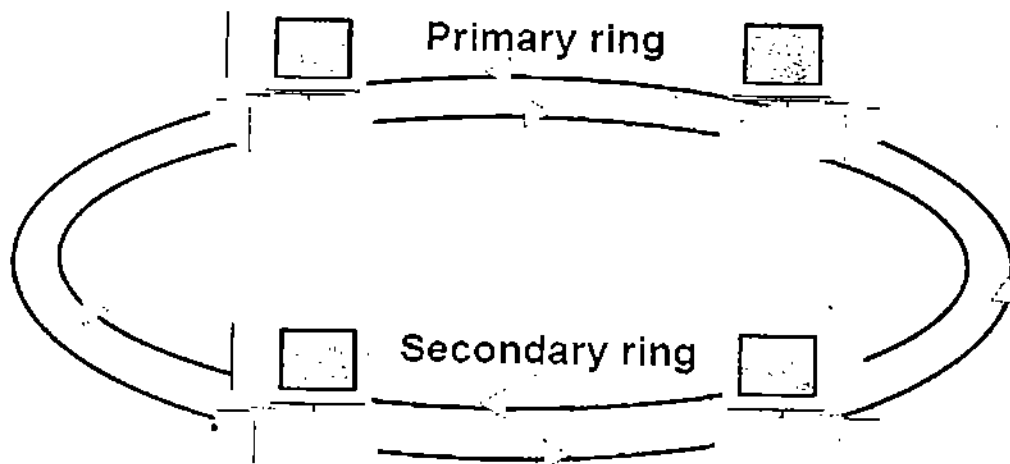
When a free token is available, a computer device is allowed to transmit its data. Unlike Token Ring, which can transmit only one frame at a time, with FDDI the computer device can transmit as many frames as possible until a predefined time limit is reached. When the device either has no more frames to send or the time limit is reached, it stops transmitting data and immediately releases an available token. In Token Ring, the free token was not released until after the sending device received the frame it sent, indicating receipt and/or errors.

Because the token is transmitted as soon as a computer device is finished transmitting frames, it is possible for a device on the ring to transmit new frames while other frames are still circulating around the ring. It is therefore possible for multiple frames from multiple devices to be on the network simultaneously.

FDDI networks are commonly used to connect LANs where either EMI or distance are problems. Computer devices can be as far apart as 2 kilometers and FDDI networks can support up to 1000 devices per LAN.

FDDI Rings

FDDI Dual Rings

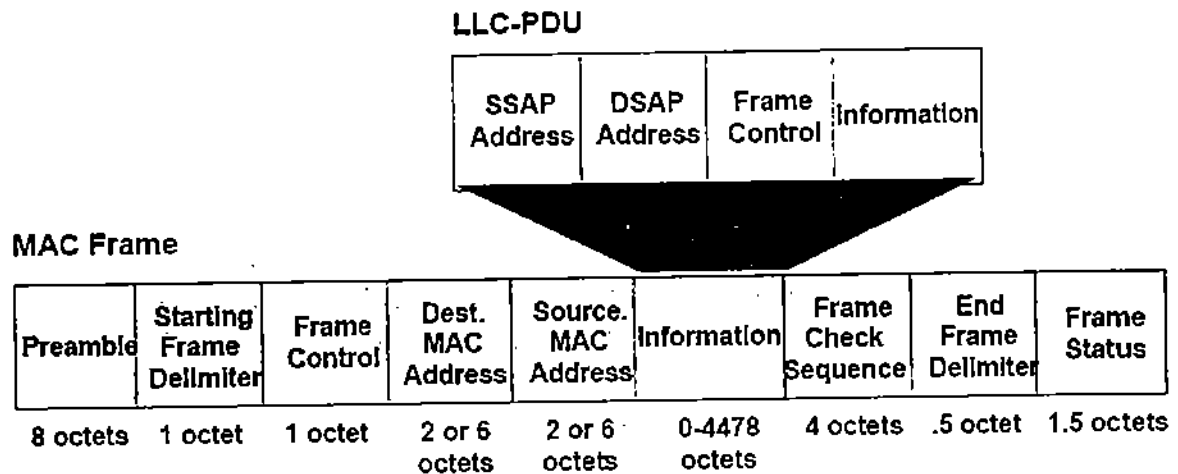


The dual rings are classified as primary or secondary. Devices on FDDI networks may be connected to both primary and secondary rings or only the primary ring. If there is a malfunction on the primary ring, the secondary ring automatically wraps to the first ring forming a complete ring once again. This is referred to as "ring wrap" and is an automatic function of FDDI. A type of claim token frame is used in FDDI networks to detect errors in transmission or timing.

A device connected to both the primary and secondary rings is classified as a dual-attached station or Class A station. A single-attached or Class B station, is attached to only the primary ring. Typically, a Class A station would be a hub and a Class B, a computer workstation. The Class B device would connect to the primary ring and through that connect to the hub.

FDDI Frame Format

In addition to the four token fields FDDI networks include the following fields as shown in the diagram below:



- Preamble: this field, which consists of all 1s, is used as a clocking mechanism and indicates the beginning of a frame.
- Start Delimiter: designates the physical start of the FDDI frame.
- Frame control: indicates the length of the address and the transmission type (synchronous/asynchronous and data/network information).
- Destination Address field.
- Source Address field.
- Data: application and upper layer protocol data.
- CRC: error checking frame.
- End Delimiter: signifies the end of the data frame.
- Frame Status field: offers further error checking.

FDDI Media

Cabling

Fiber optic cable is specified as the standard for FDDI and may be single or multi-mode fiber. Single mode fiber optic cable uses laser technology and

can transmit up to a maximum of 20 km without using repeaters. Multi-mode fiber is less expensive. Instead of laser technology, it employs LED, light emitting diode technology, and can transmit signals of to a maximum distance of 2 km.

Dual Attachment Concentrator (DAC)

A dual attachment concentrator is a hub used in FDDI networks that are capable of attaching to dual counter rotating rings. DACs have additional ports, called trunk ports, used to connect to additional concentrators or other FDDI computer devices.

FDDI Network Adapter Cards

The transceiver in FDDI networks is built into the FDDI NIC. It is a laser transceiver used to convert digital data to light impulses and vice versa.

FDDI Connectors

The connectors used in FDDI networks are Media Interface Connectors (MIC).

There are several types of MICs. They are typically color coded to assist in allowing only desirable port connections. During installation and troubleshooting when cable connections are often moved, it is important to keep the color-coded key attached to the MIC connector.

4.3 TOKEN RING CASE STUDY

In this case study, you will have the opportunity to design a Token Ring network that meets certain specifications.

The second largest video company, Blue Mood Video, is installing their first network. They have decided to install Token Ring to network their point of sale terminals for their large store. They have not decided if they need 4 or 16 Mbs.

There are 5 sales terminals, 2 servers, and 2 printers. They want to connect to the Internet. STP cable was left installed by the last occupant.

You have been hired to recommend the network speed and to design the network. Prepare a summary that addresses all of the network specifications and design questions. Include a one-page summary of the reasons for your choice and a diagram of the proposed network. Add this information to your portfolio.

Network Specifications

- The sales terminals need to respond quickly to handle the high

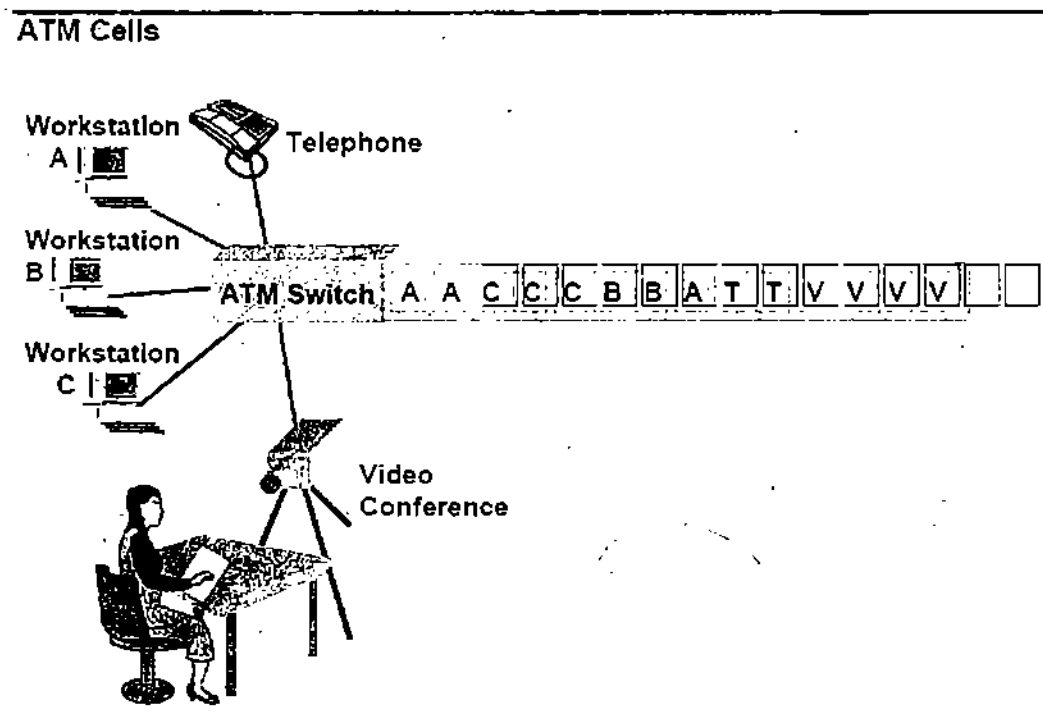
volume of sales.

- They would like to use the currently installed cable, but are not sure if it will meet their needs.
- The network needs to last and not need to be replaced in the near future.
- Add the servers and printers to appropriate locations within the store.
- List all of the necessary equipment.

Create a diagram to show your network and all of the devices. Label devices and give them MAC addresses.

4.4 ATM

One of the newest, most exciting technologies for data transmission is Asynchronous Transfer Mode (ATM), which is similar to SMDS in that it converts packets into small cells that transmit at high speeds through the network. ATM cells can transmit at speeds of 155 to 622 Mbps. This is fast enough for use on a LAN.



ATM cells travel on their own regular schedule rather than just when data needs to be transmitted. Think of ATM cells as cars on a train. When information (which could be data or video or sound) enters the ATM link, it's as if it goes to the station (the ATM switch) and boards the train (the ATM cells). The ATM switch takes packets from the LAN and puts them into the ATM cells as the cells pass by. The ATM train has an endless number of cars and moves extremely fast, so there s

never a wait to board the train. If the LAN packet is larger than a single ATM cell, then it continues filling the next cell and the next until it is completely sent. If the LAN packet only partially fills an ATM cell, then padding (meaningless data) is added to fill the cell. Because all the cells are the same size and because there is a steady stream of cells, ATM is very predictable. The protocols that handle ATM only have to deal with one size of cell and a constant speed of information. This allows these protocols to be simpler and work faster.

Another advantage to ATM is that it keeps cells in the correct order so they don't need to be sorted out at the receiving switch. One drawback is that ATM requires special switches.

ATM operates at the physical, data link, and network layers of the OSI model and can use a variety of physical transmission media. Of course, the faster transmission speeds the media can handle, the faster the data will transmit. Optical fiber (using SONET) is the best choice for the physical layer under ATM and SMDS.

4.5 Connectivity Devices

Connectivity devices are those devices used to make physical network connections. Connectivity devices operate at the physical layer of the Open Systems Interconnection Reference Model (OSI) model. The OSI model describes how computer services and procedures are standardized. This standardization allows computers to share information and enables the interconnection of various networking connectivity devices regardless of vendor.

The OSI model uses the concept of seven stacked layers to define a network communications system.

The lower three layers: Physical, Data Link, and Network, deal mostly with network-dependent (hardware) functions. The upper four layers: Transport, Session, Presentation, and Application, deal more with application-oriented (user or software) functions.

Generally speaking, these connectivity devices are passive in their actions, meaning that they do not make changes to either the data or the transmission route of that data. These devices simply provide the physical link between workstations and other network devices.

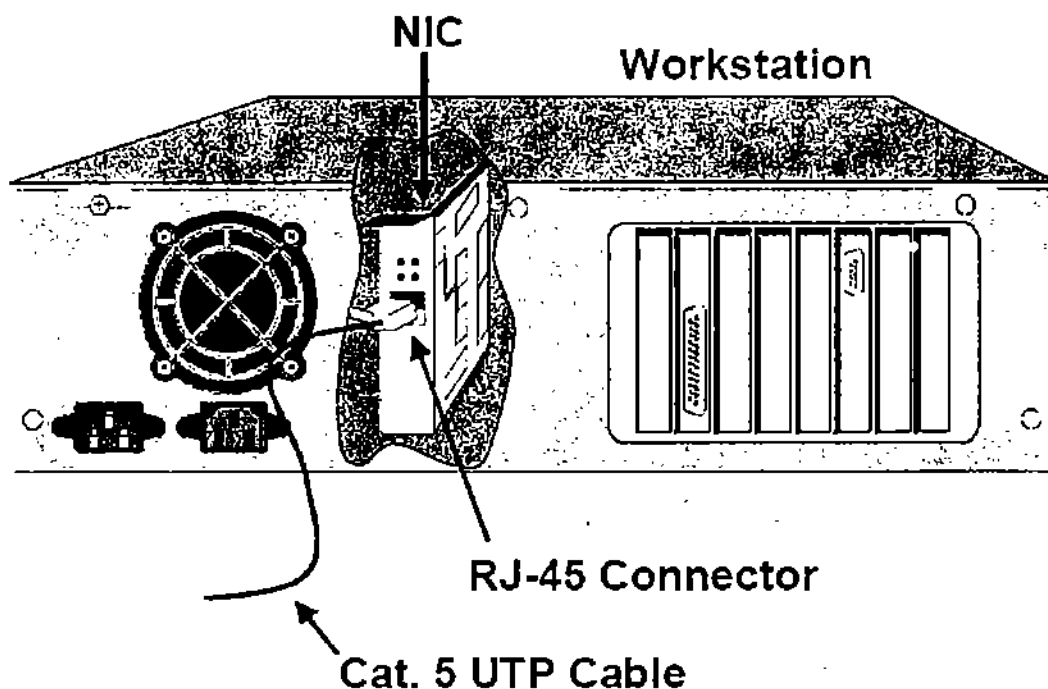
Network Interface Card

Network interface cards (NICs) are printed circuit boards that are installed in workstations. They provide the physical connection between the network cable

and the workstation. In addition, they possess the circuitry necessary to gain access to the network.

In the past, computers did not come with a pre-installed NIC. However, most modern computers are shipped with NIC already installed and ready to be connected to a network.

Network Interface Cards are Installed Inside Workstations



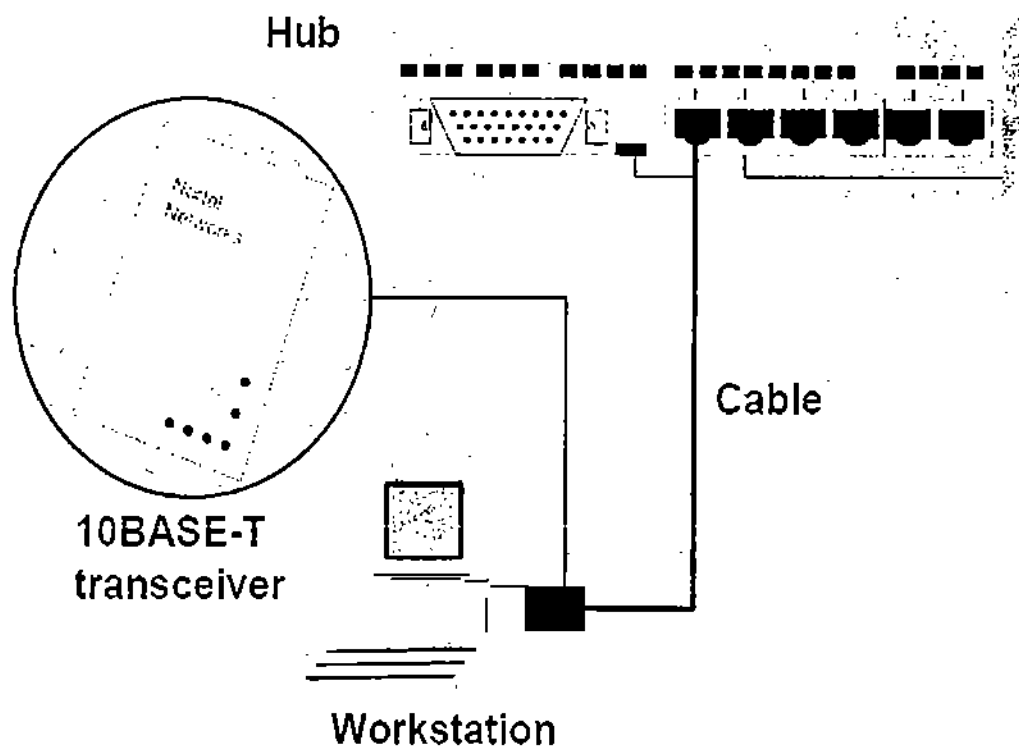
- Each NIC has a distinctive 6-byte hardware address that identifies the workstation within a segment, such as Ethernet 00-00-A2-9A-4E-10.
- Each NIC has at least one connection port where the network cable is attached. Only one port may be used at any one time.

- Although the NIC operates at the physical layer of the OSI model, it is also considered a data link layer device. Part of the NIC's function is to format information between the workstation and the network, and also, to control the transmission of data onto the wire.

4.6 Transceivers

The name transceiver is derived from the combination of the words *transmitter* and *receiver*. It is a device that both transmits and receives signals along the network cable and connects a computer to the network.

Transceivers May Be External or Located on the NIC

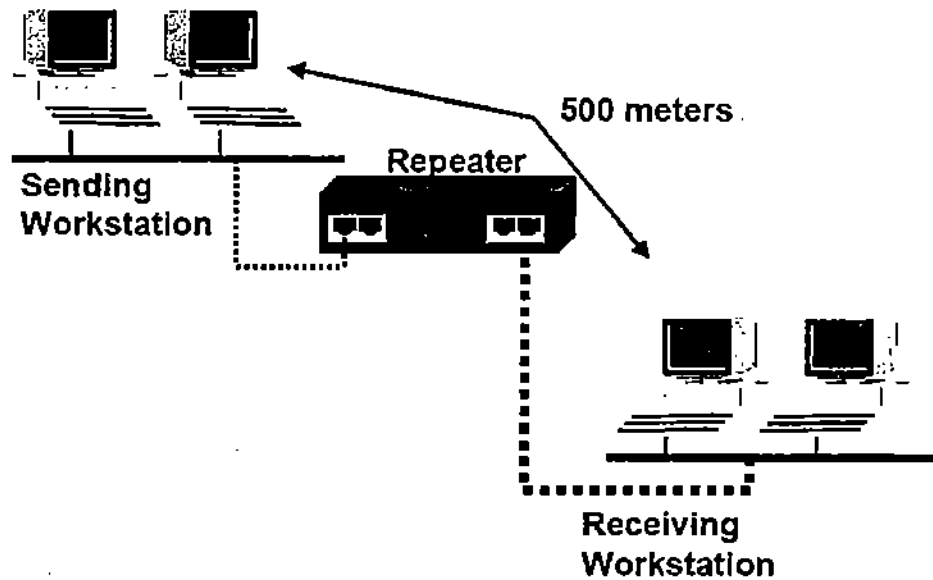


- Today, most transceivers are built into the network interface cards.
- The transceiver does not make changes to information transmitted across the network, it adapts the signals so devices connected by varying media can interpret them.
- A transceiver operates on the physical layer of the OSI model.

4.7 Repeaters

As data travels through cabling systems, a certain amount of electrical interference and signal loss is inevitable. As the need for larger networks that span greater distances developed, a solution was needed to resolve signal loss over the network. Repeaters were created to regenerate and amplify weak signals, thus extending the length of the network. The basic function of a repeater is to retime, reshape, and reamplify the data signal to its original level.

Repeaters Amplify Weak Signals

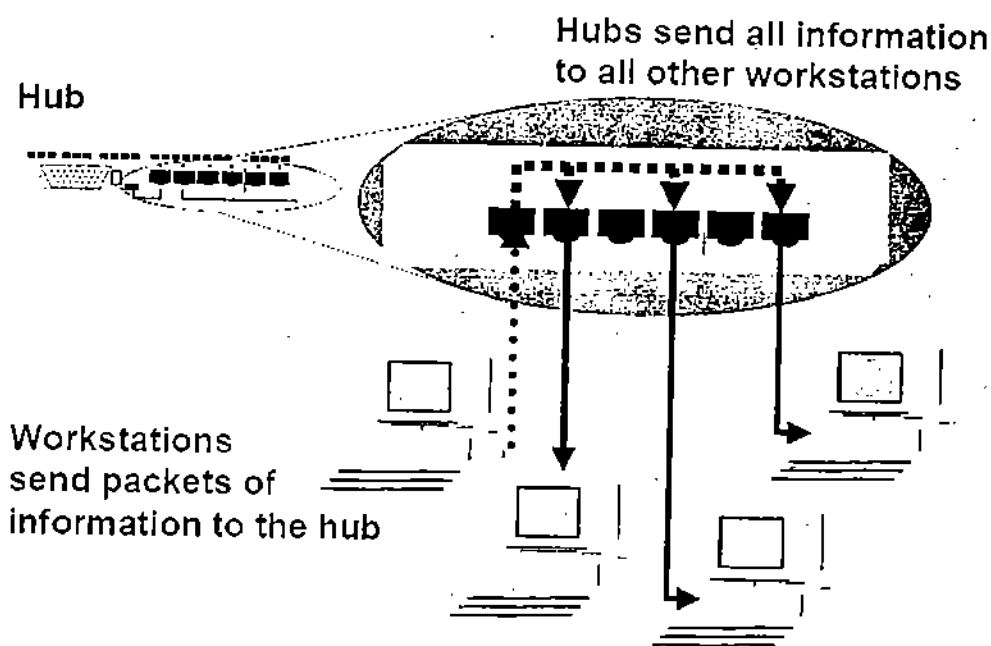


- Repeaters perform no other action on the data.
- Repeaters were originally separate devices. Today, a repeater may be a separate device or it may be incorporated into a hub.
- Repeaters operate at the physical layer of the OSI model.

4.8 Hubs

Hubs, sometimes called concentrators, reside in the core of the LAN cabling system. They are basically multiport repeaters. The hub connects workstations and sends every transmission to all the connected workstations. They work much like the old telephone party lines, where only one computer can "talk" at a time.

Hubs Connect Workstations Together



- The receptacles on the front of the hub are called ports. There are usually from 4-34 ports on a hub depending upon size of network.
- Some hubs have an additional interface port that connects to another hub, thus increasing the size of the network.
- Hub operates at the physical layer of the OSI model.

4.9 Media Dependent Adapter

A media dependent adapter (MDA), sometimes referred to as transceivers, is a module added to a hub to provide flexible cable connectivity. These optional devices provide a connection port for a cable connector that is different than the hub's ports. For example, a MDA can be used to connect a fiber-optic cable to a hub with RJ-45 connectors otherwise not compatible with fiber-optic connectors.

- Media dependent adapters are also used on servers, switches, routers, transceivers, and to make direct connections to end stations.
- MDAs function at the physical layer of the OSI model.

Media Filters

A media filter is a very specialized kind of adapter used on token ring networks. When token ring networks were first implemented, they used shielded-twisted pair (STP) cabling. Token ring networks now use unshielded-twisted pair (UTP) cables.

There is a difference in the electrical characteristics between these cabling systems. If a STP-specific network interface card in a token ring network is connected to a UTP cable, the network data signal will be damaged. Attaching media filters between the NIC and the UTP cable reduces the interference and prevents damage to the data. Media filters function at the physical layer of the OSI model.

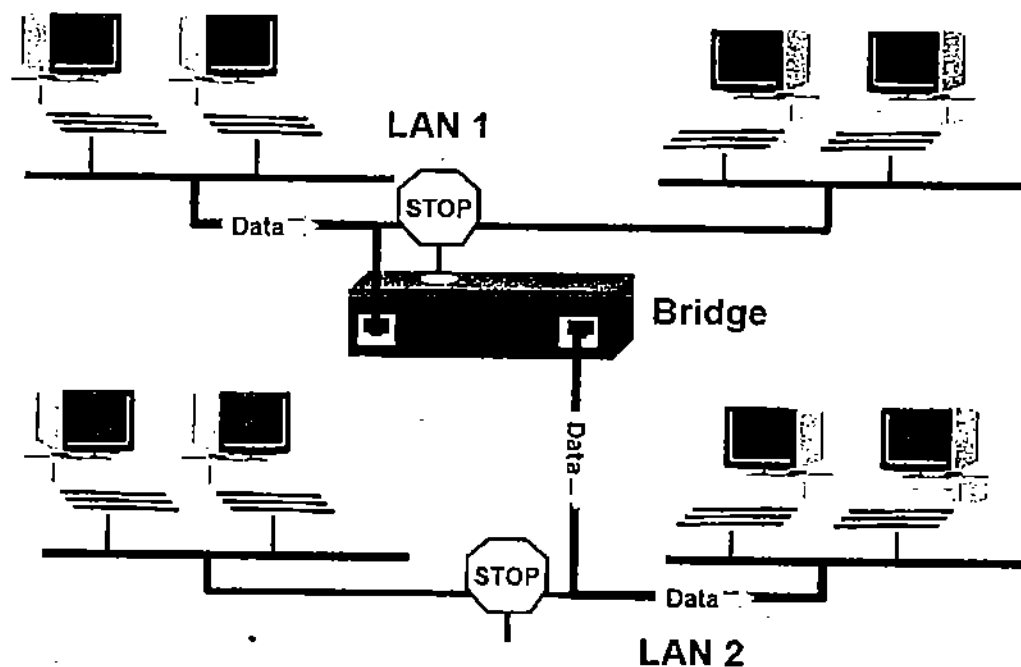
4.10 Internetworking Devices

As networks became increasingly complex, the need for internetworking devices also increased. Internetworking devices are active components rather than passive. They are considered active because they do more than simply pass data across a network. They make "intelligent" decisions and may interpret, reformat, and/or direct data as it passes through a network. Internetworking devices typically operate at OSI model layers other than the physical layer.

Bridges

Bridges connect network segments typically using the same communication protocol, passing information from one network to the other. A bridge may divide an overloaded network into smaller, more efficient networks. Bridges break networks into separate segments and direct transmission to the appropriate segment much like a police officer directs automobile traffic.

Bridges Segment Networks and Direct Traffic



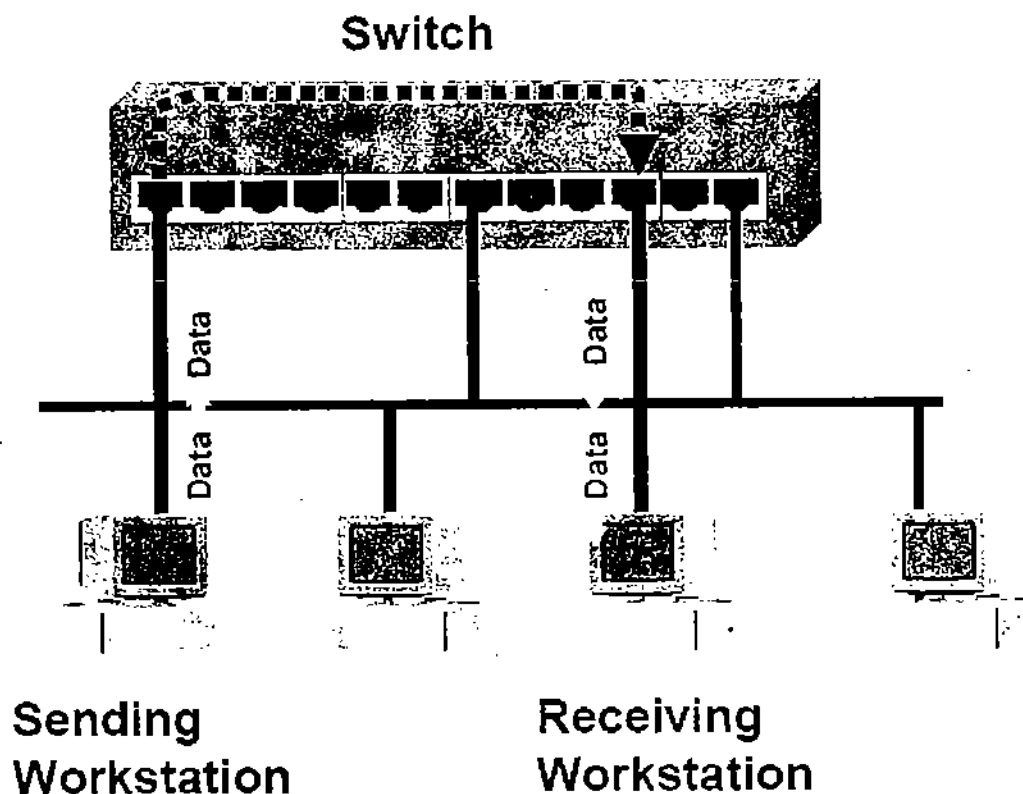
- Bridges learn which workstations are on what network segment by looking at the hardware address in the frames it receives and entering this information into a table.

- Bridge functions at data link layer of OSI model. A bridge monitors information passing over a network segment and restricts the flow of unnecessary information. It also listens to all traffic on the segment, determines the destination address, looks up the destination address in the table, and then passes the frame to the correct segment.
- Bridges can be used to connect different types of cabling from one LAN to another.

Switches

One way of relieving network congestion is to use a switch, either in place of a hub or bridge, or in addition to a hub.

Switches Direct Information to Specific Workstations



- Switches allow each workstation to transmit information over the network independent of the other workstations.
- These transmissions between computers may be simultaneous.

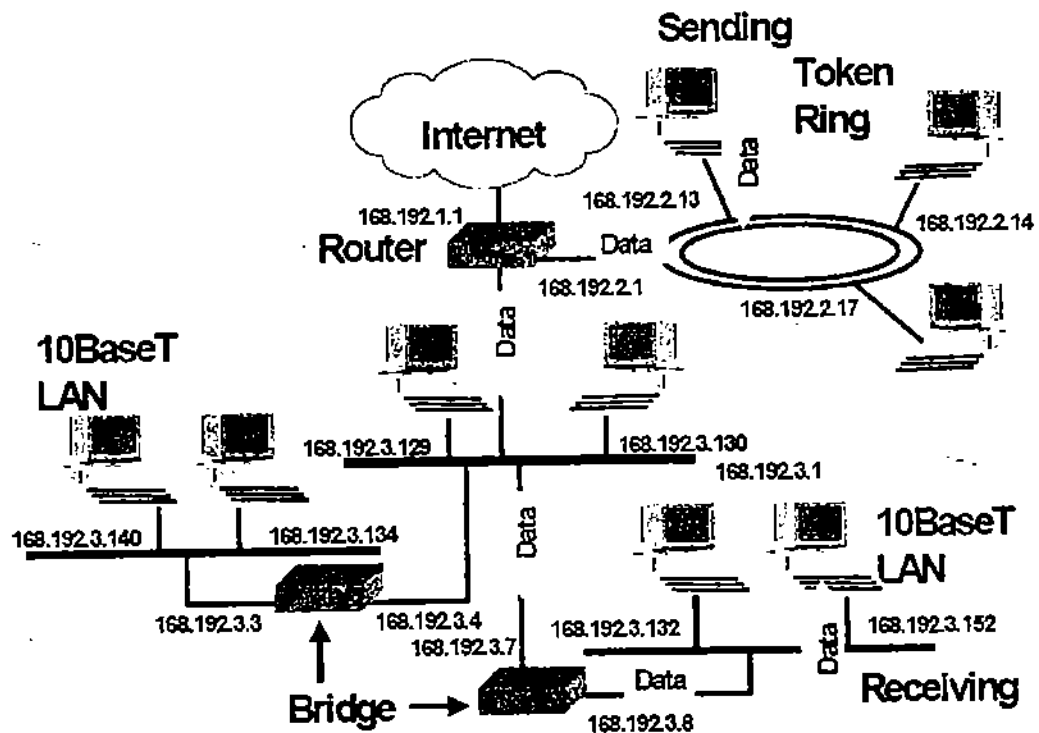
As opposed to hubs, switches are more like the modern telephone system where several private conversations take place at one time.

- Some switches can operate at different speeds, allowing the connection of cables using different bandwidths (example, 10 Mps and 100 Mps).
- There are different types of switches. Layer 3 switches function at the network layer because they can perform translation activities between network segments. Layer 2 switches function at the data link layer because they primarily direct traffic to specific workstations.

Routers

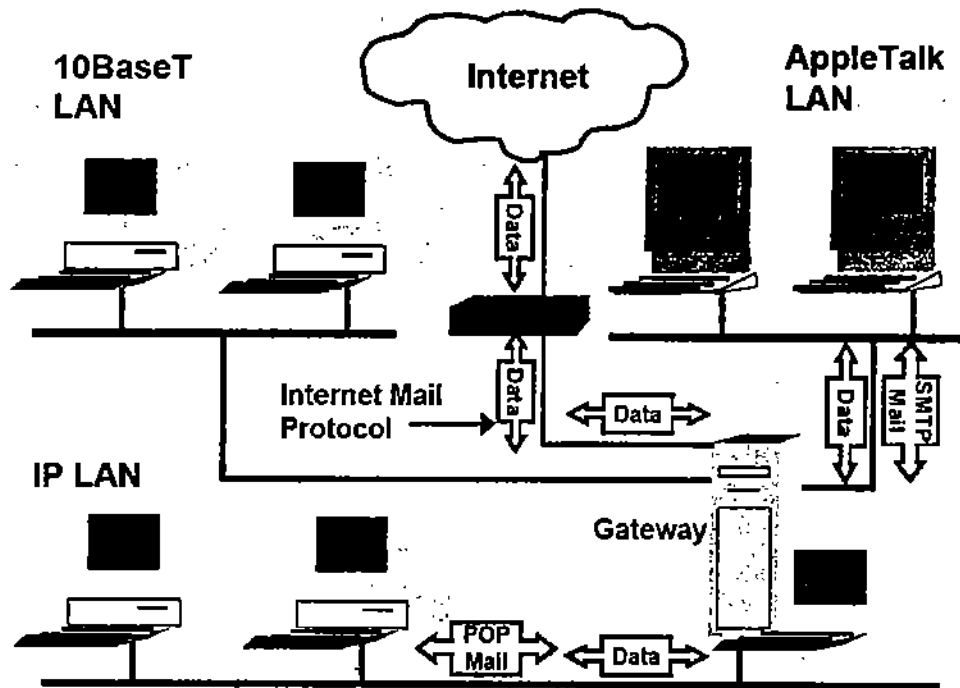
Routers link two or more different networks together, such as an Internet Protocol network. These networks can consist of various types of LAN segments, for example, Ethernet, token ring, or Fiber Distributed Data Interface (FDDI). A router receives packets and selects the optimum path to forward the packet across the network. Routers build a table of all the device addresses (routing table) across the networks. Using this table, the router forwards a transmission from the sending station to the receiving station across the best path.

Routers Direct Traffic from Different Networks via Optimum Paths



- Bridges know the 6-byte hardware address of all workstations on the network segment. In contrast, routers maintain a table of all learned network addresses, for example, 168.192.1.0, 168.192.2.0, and 168.192.3.0 are three IP network addresses.
- Routers can connect different network types such as Ethernet, token ring, and FDDI.
- Routers operate at the network level of the OSI model.

Gateways Translate Different Network Protocols



Gateways

Gateways are multi-purpose connection devices. They are able to convert the format of data in one computing environment to a format that is usable in another computer environment (for example, AppleTalk and DECnet).

The term gateway is sometimes used when referring to a router. For the purpose of this lesson, gateways are devices that link different network types and protocols. For example, gateways translate different electronic mail protocols and convey email across the Internet. Gateways can operate at all layers of OSI model since they:

- Can provide a physical link between networks.
- Create junctions between dissimilar networks.
- Translate different network protocols and/ or applications (for example, electronic mail between the Internet and a commercial online service with its own mail protocol).

4.11 Summary

In this unit we learned the following:

- Describe the characteristics of Token Ring architecture.
- Describe the characteristics of FDDI architecture.
- Design a Token Ring network to specifications.
- Describe the characteristics of ATM.
- The identification of network connectivity and internetworking devices.
- To match the appropriate level within the OSI model to which each device is assigned.
- The comparison of the functionality between connectivity and internetworking devices.

4.12 Review Questions

Part A

1. Token Ring networks operate at
 - a. 10, 100 or 1000 Mbps
 - b. 4, 16, 100 or 128 Mbps

- c. 1 Gbps
 - d. 10, 16 or 1000 Mbps
 - e. 4 or 10 Mbps
2. A device may use the token frame to send data if
- a. The Start Delimiter is 0
 - b. The End Delimiter is 1
 - c. The hardware addresses match
 - d. The Access Control signal is 0
 - e. The Start Delimiter is 1
3. The Token Ring token frame size is
- a. 16 bits
 - b. 24 bits
 - c. 32 bits
 - d. 48 bits
 - e. 64 bits
4. After copying data from the token frame, a receiving device
- a. Releases the token
 - b. Reads error control information from the frame
 - c. Uses the token frame to send new data
 - d. Destroys the token frame
 - e. Send the frame back to the sending device
5. Token Ring networks
- a. Fail more often than Ethernet networks
 - b. Fail more gracefully than Ethernet networks

- c. Never fail
 - d. Cannot accept priority assignments
 - e. Must use fiber optic cabling
6. To detect and correct errors
- a. Every device on a Token Ring network is an active monitor
 - b. Each device on a Token Ring network diagnoses its own errors
 - c. Each device gives control to its nearest active upstream neighbor
 - d. One device is designated as the active monitor
 - e. All devices send out a Claim Token
7. A MAU
- a. Connects Token Ring devices into a ring
 - b. Turn off the ring when any connected device fails
 - c. Route power to connected devices
 - d. Switch traffic to neighboring networks

Part B

1. FDDI operates at
- a. 10 Mbps
 - b. 4 or 16 Mbps
 - c. 100 Mbps
 - d. 1000 Mbps
 - e. 4.5 Mbps

2. To transport streaming media data, FDDI
 - a. Opens a switched circuit
 - b. Can synchronize data with a clocking field
 - c. Can only transmit one frame when the token has been captured
 - d. Uses fiber optic cabling

3. When an FDDI sending device has captured a token
 - a. It can only transmit one frame
 - b. It can transmit as many frames as it wants
 - c. It can only transmit error diagnostics
 - d. It **can** transmit frames until a time limit expires
 - e. It releases the token

4. A secondary ring in FDDI
 - a. Automatically steps in if the primary ring fails
 - b. Allows priority communication to select devices
 - c. Doubles communication speed
 - d. Connects to Token Ring networks

Part C

1. Which computer manufacturer uses LocalTalk?
 - a. IBM
 - b. Digital Equipment Corporation
 - c. Intel Corporation
 - d. Motorola
 - e. Apple Computer

2. LocalTalk operates at
 - a. 230 Kbps
 - b. 1 Mbps
 - c. 10 Mbps
 - d. 100 Mbps
 - e. 1 Gbps

3. LocalTalk Node Id's are assigned
 - a. By the manufacturer
 - b. By the user
 - c. By the computer
 - d. By the server

4. LocalTalk connects using
 - a. Fiber optic cable
 - b. Coaxial cable
 - c. Twisted pair cable
 - d. Microwave radio
 - e. Photonic switches

Part D

1. ARCNet operates at
 - a. 230 Kbps
 - b. 1 Mbps
 - c. 2.5 Mbps
 - d. 10 Mbps
 - e. 1 Gbps

2. ARCNet tokens are passed by
 - a. nearest active upstream neighbor
 - b. nearest active downstream neighbor
 - c. Random walk
 - d. MAC address
 - e. NAUN computer device

3. ARCNet connects using
 - a. Twisted pair or coaxial cable
 - b. Fiber optic cable
 - c. Infrared light
 - d. Microwave radio

Connectivity and Internetworking Devices

Part A

Identify each item as either a connectivity device or an internetworking device by placing a "C" or an "I" next to the item.

	1. Bridge		7. Switch
	2. Multistation Access Unit		8. Gateway
	3. Transceiver		9. Router
	4. Network Interface Card		10. Repeater
	5. Hub		11. Media Filter
	6. Media Dependent Adapter		

Part B

Identify which layer of the OSI model each device is assigned.

	1. Bridge
	2. Repeater
	3. Network Interface Card
	4. Router
	5. Transceiver
	6. Multistation Access Unit
	7. Hub
	8. Switch
	9. Media Dependent Adapter
	10. Media Filter
	11. Gateway



Uttar Pradesh
Rajarshi Tandon Open University

BLISS-03

**Computer Network and
Security Maintenance**

BLOCK

3

Internet Connectivity

Unit 1

The Internet

5

Unit 2

The Internet Services

16

Unit 3

ISDN and Bridge-Routers

27

Unit 4

ISP Connectivity

43

Block Introduction:

This is the third block of the Data Communication and Networking. In this block the Internet and its services are described. Next, ISDN and Bridge-routers are briefly introduced. The block also covers the aspects of ISP connectivity.

In unit 1, there is introduction to the Internet. This unit covers Internet usage and architecture of the Internet. Main focus is given to the TCP/IP reference model and unified networks.

Unit 2 starts by listing the Internet services. It describes E mail, file transfer, telnet and lease lines.

In unit 3 ISDN and Bridge-routers are introduced. The unit defines ISDN and provides the history and advantages of ISDN. Next, ISDN interfaces and protocols are described. Lastly, the functioning of Bridge-routers is explained.

Unit 4 is mainly focused about ISP connectivity. The unit starts by giving introduction to Internet Service Providers. Then it explains Cable MODEM, DSL, SHDSL, and how ISP connect to the Internet. Finally, Broadband access is described briefly.

Unit I – The Internet

What You Will Learn (Objectives)

After completing this unit you will be able to:

- Understand and define Internet.
- Internet usage and architecture of the Internet
- Describe Internet connectivity rules, Internet Protocols
- Learn TCP/IP Reference Model and unified networks

1.1 INTRODUCTION

The subject of computer networking covers many different kinds of networks, large and small, well known and less well known. They have different goals, scales, and technologies. In the following sections, we will look at some examples, to get an idea of the variety one finds in the area of computer networking. We will start with the Internet, probably the best known network, and look at its history, evolution, and technology.

The Internet is not a network at all, but a vast collection of different networks that use certain common protocols and provide certain common services. It is an unusual system in that it was not planned by anyone and is not controlled by anyone. To better understand it, let us start from the beginning and see how it has developed and why.

The Internet

The Internet is the world's largest network, spanning the globe. Its origins date back to the 1960's. It consists of a super network of interconnected government, education, and business networks.

In the 1960's the U.S. Department of Defense (DOD) wanted a network that would allow communications throughout the United States even if the existing telecommunications lines were destroyed by enemy attack. Funded through the DOD's Advanced Research Projects Agency, a new network was formed called ARPANet with only four computers, three located in California and one in Nevada. ARPANet was limited by the fact that different systems could not connect to each other. With funding from military sources, the universities and government contractors continued to improve on ARPANet.

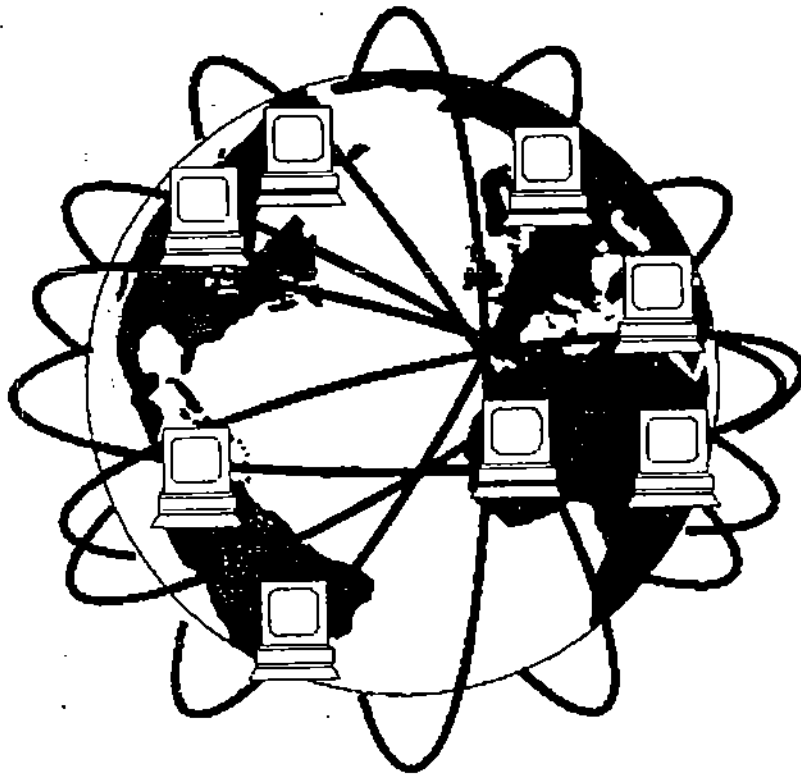
- 1973—Transmission Control Protocol (TCP) and Internet Protocol (IP) emerged, allowing various systems to interconnect.
- 1975—The Defense Communications Agency became responsible for ARPANet.
- 1979—Academic and research facilities were allowed to connect to ARPANet leading to a rapid growth of the network.
- 1980—Tim Berners-Lee, a software engineer, developed HTML (HyperText Mark-up Language) used in the World Wide Web.
- 1986—The National Science Foundation initiated their own high-speed network called NSFNet that connected various regional networks. The term "Internet" was adopted.

- 1988—MCIMail was given permission to connect to the Internet.
- 1989 – The graphical interface for the Internet called the World Wide Web (WWW) was proposed using HTML. CompuServe, ATTMail, and Sprintmail connected to the Internet. ARPANet was dismantled.
- 1991—Tim Berners-Lee developed the HyperText Transfer Protocol (HTTP) and introduced the WWW on the Internet.
- 1993—NSFNet began to be dismantled.
- 1999—Over 150 million people use the Internet/WWW each week.

Today the Internet is the largest interconnection of networks in the world. The Internet allows communication and sharing of information between users all over the world through the use of connected computers and an extremely efficient technology called packet switching which moves huge quantities of data around the world. It has grown to thousands of interconnected networks with millions of users worldwide.

With the development of a graphical user interface for the World Wide Web, more people can access the web with less training. The demand for access has led to thousands of web servers linking the world to retail stores, on-line databases, research resources, government publications, public libraries and more.

The Internet: World's Largest Network



1.2 Internet Usage

The number of networks, machines, and users connected to the ARPANET grew rapidly after TCP/IP became the only official protocol on January 1, 1983. When NSFNET and the ARPANET were interconnected, the growth became exponential. Many regional networks joined up, and connections were made to networks in Canada, Europe, and the Pacific.

Sometime in the mid-1980s, people began viewing the collection of networks as an Internet, and later as the Internet, although there was no official dedication with some politician breaking a bottle of champagne over a fuzz ball.

The glue that holds the Internet together is the TCP/IP reference model and TCP/IP protocol stack. TCP/IP makes universal service possible and can be compared to the adoption of standard gauge by the railroads in the 19th century or the adoption of common signaling protocols by all the telephone companies.

What does it actually mean to be on the Internet? Our definition is that a machine is on the Internet if it runs the TCP/IP protocol stack, has an IP address, and can send IP packets to all the other machines on the Internet. The mere ability to send and receive electronic mail is not enough, since e-mail is gatewayed to many networks outside the Internet. However, the issue is clouded somewhat by the fact that millions of personal computers can call up an Internet service provider using a modem, be assigned a temporary IP address, and send IP packets to other Internet hosts. It makes sense to regard such machines as being on the Internet for as long as they are connected to the service provider's router.

Traditionally (meaning 1970 to about 1990), the Internet and its predecessors had four main applications:

1. **E-mail.** The ability to compose, send, and receive electronic mail has been around since the early days of the ARPANET and is enormously popular. Many people get dozens of messages a day and consider it their primary way of interacting with the outside world, far outdistancing the telephone and snail mail. E-mail programs are available on virtually every kind of computer these days.
2. **News.** Newsgroups are specialized forums in which users with a common interest can exchange messages. Thousands of newsgroups exist, devoted to technical and nontechnical topics, including computers, science, recreation, and politics. Each newsgroup has its own etiquette, style, and customs, and woe beside anyone violating them.
3. **Remote login.** Using the telnet, rlogin, or ssh programs, users anywhere on the Internet can log on to any other machine on which they have an account.
4. **File transfer.** Using the FTP program, users can copy files from one machine on the Internet to another. Vast numbers of articles, databases, and other information are available this way.

Up until the early 1990s, the Internet was largely populated by academic, government, and industrial researchers. One new application, the WWW (World Wide Web) changed all that and brought millions of new, nonacademic users to the net. This application, invented by CERN physicist Tim Berners-Lee, did not change any of the underlying facilities but made them easier to use. Together with the Mosaic browser, written by Marc Andreessen at the National Center

for Supercomputer Applications in Urbana, Illinois, the WWW made it possible for a site to set up a number of pages of information containing text, pictures, sound, and even video, with embedded links to other pages. By clicking on a link, the user is suddenly transported to the page pointed to by that link. For example, many companies have a home page with entries pointing to other pages for product information, price lists, sales, technical support, communication with employees, stockholder information, and more.

Numerous other kinds of pages have come into existence in a very short time, including maps, stock market tables, library card catalogs, recorded radio programs, and even a page pointing to the complete text of many books whose copyrights have expired (Mark Twain, Charles Dickens, etc.). Many people also have personal pages (home pages).

Much of this growth during the 1990s was fueled by companies called ISPs (Internet Service Providers). These are companies that offer individual users at home the ability to call up one of their machines and connect to the Internet, thus gaining access to e-mail, the WWW, and other Internet services. These companies signed up tens of millions of new users a year during the late 1990s, completely changing the character of the network from an academic and military playground to a public utility, much like the telephone system. The number of Internet users now is unknown, but is certainly hundreds of millions worldwide and will probably hit 1 billion fairly soon.

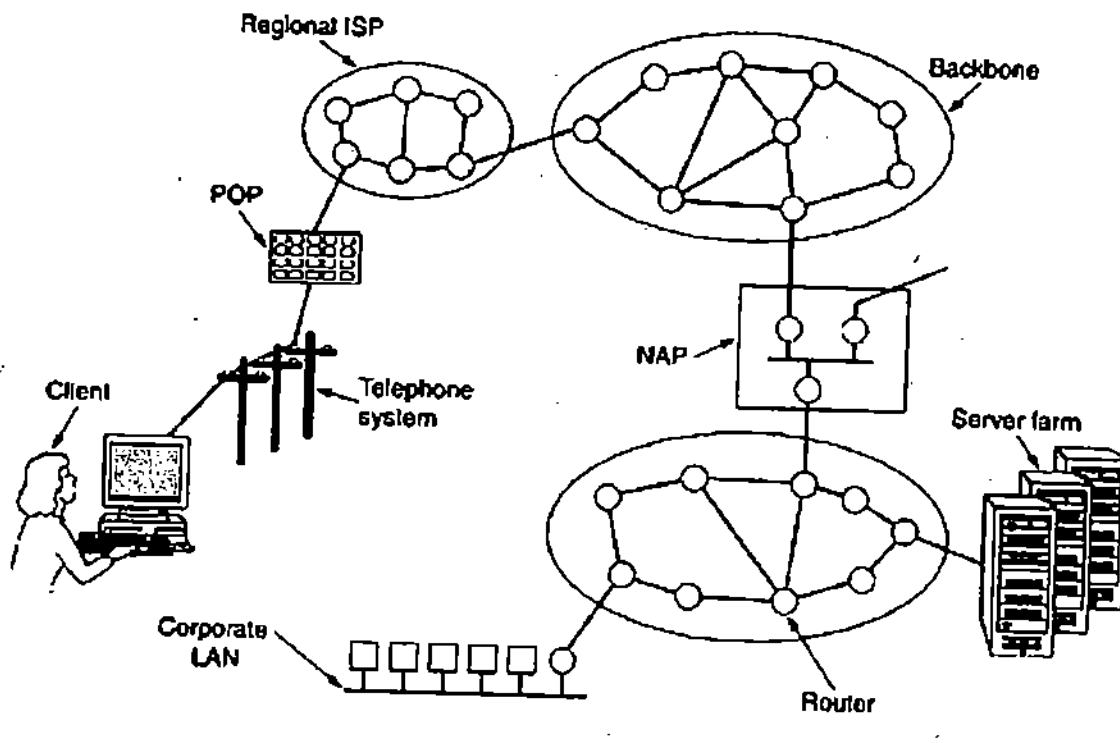
1.3 Architecture of the Internet

In this section we will attempt to give a brief overview of the Internet today. Due to the many mergers between telephone companies (telcos) and ISPs, the waters have become muddled and it is often hard to tell who is doing what. Consequently, this description will be of necessity somewhat simpler than reality. The big picture is shown in Figure. Let us examine this figure piece by piece now.

A good place to start is with a client at home. Let us assume our client calls his or her ISP over a dial-up telephone line, as shown in the figure. The modem is a card within the PC that converts the digital signals the computer produces to analog signals that can pass unhindered over the telephone system. These signals are transferred to the ISP's POP (Point of Presence), where they are removed from the telephone system and injected into the ISP's regional network. From this point on, the system is fully digital and packet switched. If the ISP is the local telco, the POP will probably be located in the telephone switching office where the telephone wire from the client terminates. If the ISP is not the local telco, the POP may be a few switching offices down the road.

The ISP's regional network consists of interconnected routers in the various cities the ISP serves. If the packet is destined for a host served directly by the ISP, the packet is delivered to the host. Otherwise, it is handed over to the ISP's backbone operator.

Figure: Overview of Internet



At the top of the food chain are the major backbone operators, companies like AT&T and Sprint. They operate large international backbone networks, with thousands of routers connected by high-bandwidth fiber optics. Large corporations and hosting services that run server farms (machines that can serve thousands of Web pages per second) often connect directly to the backbone. Backbone operators encourage this direct connection by renting space in what are called carrier hotels, basically equipment racks in the same room as the router to allow short, fast connections between server farms and the backbone.

If a packet given to the backbone is destined for an ISP or company served by the backbone, it is sent to the closest router and handed off there. However, many backbones, of varying sizes, exist in the world, so a packet may have to go to a competing backbone. To allow packets to hop between backbones, all the major backbones connect at the NAPs discussed earlier. Basically, a NAP is a room full of routers, at least one per backbone. A LAN in the room connects all the routers, so packets can be forwarded from any backbone to any other backbone. In addition to being interconnected at NAPs, the larger backbones have numerous direct connections between their routers, a technique known as private peering. One of the many paradoxes of the Internet is that ISPs who publicly compete with one another for customers often privately cooperate to do private peering.

This ends our quick tour of the Internet. We will have a great deal to say about the individual components and their design, algorithms, and protocols in subsequent chapters. Also worth mentioning in passing is that some companies have interconnected all their existing internal

networks, often using the same technology as the Internet. These intranets are typically accessible only within the company but otherwise work the same way as the Internet.

1.4 The Internet Protocol – TCP/IP Reference Model

Let us now turn from the OSI reference model to the reference model used in the grandparent of all wide area computer networks, the ARPANET, and its successor, the worldwide Internet. The ARPANET was a research network sponsored by the DoD (U.S. Department of Defense). It eventually connected hundreds of universities and government installations, using leased telephone lines. When satellite and radio networks were added later, the existing protocols had trouble interworking with them, so a new reference architecture was needed. Thus, the ability to connect multiple networks in a seamless way was one of the major design goals from the very beginning. This architecture later became known as the **TCP/IP Reference Model**, after its two primary protocols. It was first defined by Cerf and Kahn, 1974.

Given the DoD's worry that some of its precious hosts, routers, and internetwork gateways might get blown to pieces at a moment's notice, another major goal was that the network be able to survive loss of subnet hardware, with existing conversations not being broken off. In other words, DoD wanted connections to remain intact as long as the source and destination machines were functioning, even if some of the machines or transmission lines in between were suddenly put out of operation. Furthermore, a flexible architecture was needed since applications with divergent requirements were envisioned, ranging from transferring files to real-time speech transmission.

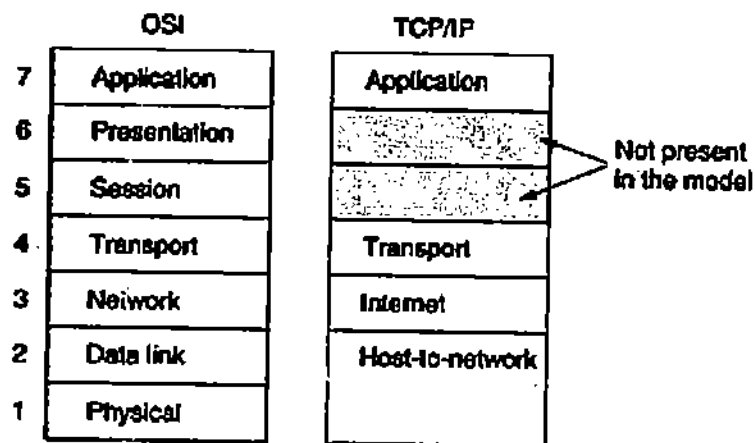
1.4.1 The Internet Layer

All these requirements led to the choice of a packet-switching network based on a connectionless internetwork layer. This layer, called the internet layer, is the linchpin that holds the whole architecture together. Its job is to permit hosts to inject packets into any network and have them travel independently to the destination (potentially on a different network). They may even arrive in a different order than they were sent, in which case it is the job of higher layers to rearrange them, if in-order delivery is desired. Note that "internet" is used here in a generic sense, even though this layer is present in the Internet.

The analogy here is with the (snail) mail system. A person can drop a sequence of international letters into a mail box in one country, and with a little luck, most of them will be delivered to the correct address in the destination country. Probably the letters will travel through one or more international mail gateways along the way, but this is transparent to the users. Furthermore, that each country (i.e., each network) has its own stamps, preferred envelope sizes, and delivery rules is hidden from the users.

The internet layer defines an official packet format and protocol called IP (Internet Protocol). The job of the internet layer is to deliver IP packets where they are supposed to go. Packet routing is clearly the major issue here, as is avoiding congestion. For these reasons, it is reasonable to say that the TCP/IP internet layer is similar in functionality to the OSI network layer. Figure shows this correspondence.

Figure: The TCP/IP Reference Model

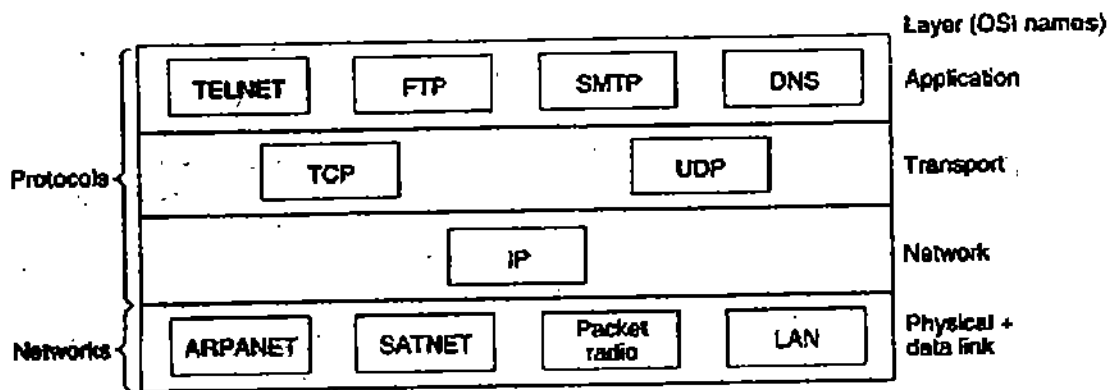


1.4.2 The Transport Layer

The layer above the internet layer in the TCP/IP model is now usually called the transport layer. It is designed to allow peer entities on the source and destination hosts to carry on a conversation, just as in the OSI transport layer. Two end-to-end transport protocols have been defined here. The first one, TCP (Transmission Control Protocol), is a reliable connection-oriented protocol that allows a byte stream originating on one machine to be delivered without error on any other machine in the internet. It fragments the incoming byte stream into discrete messages and passes each one on to the internet layer. At the destination, the receiving TCP process reassembles the received messages into the output stream. TCP also handles flow control to make sure a fast sender cannot swamp a slow receiver with more messages than it can handle.

The second protocol in this layer, UDP (User Datagram Protocol), is an unreliable, connectionless protocol for applications that do not want TCP's sequencing or flow control and wish to provide their own. It is also widely used for one-shot, client-server-type request-reply queries and applications in which prompt delivery is more important than accurate delivery, such as transmitting speech or video. The relation of IP, TCP, and UDP is shown in figure. Since the model was developed, IP has been implemented on many other networks.

Figure: Protocols and networks in the TCP/IP model initially



1.4.3 The Application Layer

The TCP/IP model does not have session or presentation layers. No need for them was perceived, so they were not included. Experience with the OSI model has proven this view correct: they are of little use to most applications.

On top of the transport layer is the application layer. It contains all the higher-level protocols. The early ones included virtual terminal (TELNET), file transfer (FTP), and electronic mail (SMTP), as shown in figure above. The virtual terminal protocol allows a user on one machine to log onto a distant machine and work there. The file transfer protocol provides a way to move data efficiently from one machine to another. Electronic mail was originally just a kind of file transfer, but later a specialized protocol (SMTP) was developed for it. Many other protocols have been added to these over the years: the Domain Name System (DNS) for mapping host names onto their network addresses, NNTP, the protocol for moving USENET news articles around, and HTTP, the protocol for fetching pages on the World Wide Web, and many others.

1.4.4 The Host-to-Network Layer

Below the internet layer is a great void. The TCP/IP reference model does not really say much about what happens here, except to point out that the host has to connect to the network using some protocol so it can send IP packets to it. This protocol is not defined and varies from host to host and network to network.

1.5 Unified Networks

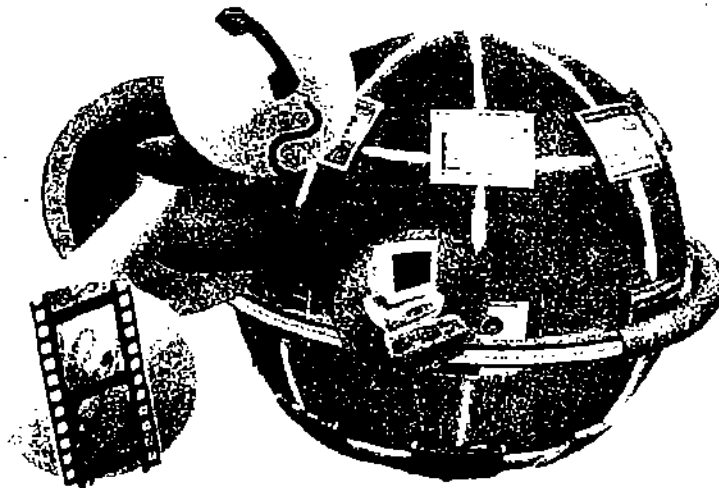
"Since the invention of the telephone, almost every civilian communications network has been based on the concept of circuits." The telephone system has been based on a system of circuit-switched networks that allow users to communicate over a dedicated line for as long as needed. The network stays open until the user disconnects, even if there are times when no data (i.e. conversation) is being transmitted. While the circuit is open, no one else can use that circuit and costs accrue even when no data is being sent.

Circuit switching is often used for the exchange of information between computer networks. The system works much the same way, the connection or *circuit* between your computer, the switch and the network remains open, costs accrue, and no one else can use the circuit while the circuit is open even when no data is transmitting.

The Internet uses another technology called packet switching, which breaks long messages into smaller, multiple packets for transmission across the most logical and efficient pathway. The packets are then reassembled at the receiving end. This allows more than one computer to transmit data over the wire at a time. Packet switching is faster and more efficient than circuit switching.

Today, these technologies are becoming unified into one infrastructure erasing the lines between computers and telephony. Unified networks bring together many different types of products and technologies, including circuit and packet; switching and routing; optical, wireline, and wireless transmission, using a single management system.

The Internet is changing how we communicate, how we relate to one another, and even how we live and do business. With unified networks, call centers over the Internet are emerging to increase the interaction between businesses and their customers. Customers will pull up a web page and click a button to call a business to place an order or request other services, without leaving the Internet session or dialing a number. Voice, data, and video will be unified into one network of efficiency and speed.



Unified Networks will revolutionize the way we communicate

1.6 Summary

This unit introduces the concept and history of Internet. Then it describes Internet architecture and its usage. Next is the discussion about Internet connectivity rulers and Internet protocols including the Internet Layer, Transport Layer, Application Layer, and the Host-to-Network Layer. Finally, we learn about TCP/IP reference model and unified networks.

1.7 Review Questions

1. Diagram a simple timeline of the history of Internet and write a summary about the milestones in Internet.
2. Describe the concept of Unified Networks and how they will improve global communications.
3. What do you mean by Internet?
4. Draw the architecture of the Internet. What are the usages of Internet?
5. Describe Internet connectivity rulers and Internet protocols.
6. What is TCP/IP reference model? Describe transport layer, application layer and the host-to-network layer.
7. The Internet is made up of a large number of networks. Their arrangement determines the topology of the Internet. A considerable amount of information about the Internet topology is available on line. Use a search engine to find out more about the Internet topology and write a short report summarizing your findings.
8. List two advantages and two disadvantages of having international standards for network protocols.
9. The Internet is roughly doubling in size every 18 months. Although no one really knows for sure, one estimate put the number of hosts on it at 100 million in 2001. Use these data to compute the expected number of Internet hosts in the year 2010. Do you believe this? Explain why or why not.
10. List two ways in which the OSI reference model and the TCP/IP reference model are the same. Now list two ways in which they differ.

1.6 Summary

This unit introduces the concept and history of Internet. Then it describes Internet architecture and its usage. Next is the discussion about Internet connectivity rulers and Internet protocols including the Internet Layer, Transport Layer, Application Layer, and the Host-to-Network Layer. Finally, we learn about TCP/IP reference model and unified networks.

1.7 Review Questions

1. Diagram a simple timeline of the history of Internet and write a summary about the milestones in Internet.
2. Describe the concept of Unified Networks and how they will improve global communications.
3. What do you mean by Internet?
4. Draw the architecture of the Internet. What are the usages of Internet?
5. Describe Internet connectivity rulers and Internet protocols.
6. What is TCP/IP reference model? Describe transport layer, application layer and the host-to-network layer.
7. The Internet is made up of a large number of networks. Their arrangement determines the topology of the Internet. A considerable amount of information about the Internet topology is available on line. Use a search engine to find out more about the Internet topology and write a short report summarizing your findings.
8. List two advantages and two disadvantages of having international standards for network protocols.
9. The Internet is roughly doubling in size every 18 months. Although no one really knows for sure, one estimate put the number of hosts on it at 100 million in 2001. Use these data to compute the expected number of Internet hosts in the year 2010. Do you believe this? Explain why or why not.
10. List two ways in which the OSI reference model and the TCP/IP reference model are the same. Now list two ways in which they differ.

Unit II – The Internet Services

What You Will Learn (Objectives)

After completing this unit you will be able to:

- Know Internet services
- Describe Email
- Describe file transfer, telnet
- Learn about lease lines

2.1 Introduction: Internet Services

Traditionally (meaning 1970 to about 1990), the Internet and its predecessors had four main applications:

1. **E-mail.** The ability to compose, send, and receive electronic mail has been around since the early days of the ARPANET and is enormously popular. Many people get dozens of messages a day and consider it their primary way of interacting with the outside world, far outdistancing the telephone and snail mail. E-mail programs are available on virtually every kind of computer these days.
2. **News.** Newsgroups are specialized forums in which users with a common interest can exchange messages. Thousands of newsgroups exist, devoted to technical and nontechnical topics, including computers, science, recreation, and politics. Each newsgroup has its own etiquette, style, and customs.
3. **Remote login.** Using the telnet, rlogin, or ssh programs, users anywhere on the Internet can log on to any other machine on which they have an account.
4. **File transfer.** Using the FTP program, users can copy files from one machine on the Internet to another. Vast numbers of articles, databases, and other information are available this way.

Up until the early 1990s, the Internet was largely populated by academic, government, and industrial researchers. One new application, the WWW (World Wide Web) changed all that and brought millions of new, nonacademic users to the net. This application, invented by CERN physicist Tim Berners-Lee, did not change any of the underlying facilities but made them easier to use. Together with the Mosaic browser, written by Marc Andreessen at the National Center for Supercomputer Applications in Urbana, Illinois, the WWW made it possible for a site to set up a number of pages of information containing text, pictures, sound, and even video, with embedded links to other pages. By clicking on a link, the user is suddenly transported to the page pointed to by that link. For example, many companies have a home page with entries pointing to other pages for product information, price lists, sales, technical support, communication with employees, stockholder information, and more.

Numerous other kinds of pages have come into existence in a very short time, including maps, stock market tables, library card catalogs, recorded radio programs, and even a page pointing to the complete text of many books whose copyrights have expired. Many people also have personal pages (home pages).

Much of this growth during the 1990s was fueled by companies called ISPs (Internet Service Providers). These are companies that offer individual users at home the ability to call up one of their machines and connect to the Internet, thus gaining access to e-mail, the WWW, and other Internet services. These companies signed up tens of millions of new users a year during the late 1990s, completely changing the character of the network from an academic and military playground to a public utility, much like the telephone system. The number of Internet users now is unknown, but is certainly hundreds of millions worldwide and will probably hit 1 billion fairly soon.

Internet Service Providers (ISPs)

These are companies that offer individual users at home the ability to call up one of their machines and connect to the Internet, thus gaining access to e-mail, the WWW, and other Internet services. These companies signed up tens of millions of new users a year during the late 1990s, completely changing the character of the network from an academic and military playground to a public utility, much like the telephone system.

2.2 Electronic Mail (e-mail)

Electronic mail, or e-mail, as it is known to its many fans, has been around for over two decades. Before 1990, it was mostly used in academia. During the 1990s, it became known to the public at large and grew exponentially to the point where the number of e-mails sent per day now is vastly more than the number of snail mail (i.e., paper) letters.

E-mail, like most other forms of communication, has its own conventions and styles. In particular, it is very informal and has a low threshold of use. People who would never dream of calling up or even writing a letter to a Very Important Person do not hesitate for a second to send a sloppily-written e-mail.

E-mail is full of jargon such as BTW (By The Way), ROTFL (Rolling On The Floor Laughing), and IMHO (In My Humble Opinion). Many people also use little ASCII symbols called smileys or emoticons in their e-mail. A few of the more interesting ones are reproduced in table. For most, rotating the book 90 degrees clockwise will make them clearer.

Table: Some smileys

Smiley	Meaning	Smiley	Meaning	Smiley	Meaning
:~)	I'm happy	=!:-)	Abe Lincoln	:+)	Big nose
:-(I'm sad/angry	=):~)	Uncle Sam	:~))	Double chin
:~	I'm apathetic	*<:~)	Santa Claus	:~(Mustache
:~)	I'm winking	<:~(Dunce	#:~)	Matted hair
:~(O)	I'm yelling	(:~	Australian	8~)	Wears glasses
:~(*)	I'm vomiting	:~)X	Man with bowtie	C:~)	Large brain

The first e-mail systems simply consisted of file transfer protocols, with the convention that the first line of each message (i.e., file) contained the recipient's address. As time went on, the limitations of this approach became more obvious.

As experience was gained, more elaborate e-mail systems were proposed. In 1982, the ARPANET e-mail proposals were published as RFC 821 (transmission protocol) and RFC 822 (message format). Minor revisions, RFC 2821 and RFC 2822, have become Internet standards, but everyone still refers to Internet e-mail as RFC 822.

3.2.1 The User Agent

E-mail systems have two basic parts, as we have seen: the user agents and the message transfer agents. In this section we will look at the user agents. A user agent is normally a program (sometimes called a mail reader) that accepts a variety of commands for composing, receiving, and replying to messages, as well as for manipulating mailboxes. Some user agents have a fancy menu- or icon-driven interface that requires a mouse, whereas others expect 1-character commands from the keyboard. Functionally, these are the same. Some systems are menu- or icon-driven but also have keyboard shortcuts.

Sending E-mail

To send an e-mail message, a user must provide the message, the destination address, and possibly some other parameters. The message can be produced with a free-standing text editor, a word processing program, or possibly with a specialized text editor built into the user agent. The destination address must be in a format that the user agent can deal with. Many user agents expect addresses of the form `user@dns-address`. Since we have studied DNS earlier in this chapter, we will not repeat that material here.

However, it is worth noting that other forms of addressing exist. In particular, X.400 addresses look radically different from DNS addresses. They are composed of attribute = value pairs separated by slashes, for example,

```
/C=US/ST=MASSACHUSETTS/L=CAMBRIDGE/PA=360 MEMORIAL DR./CN=KEN SMITH/
```

This address specifies a country, state, locality, personal address and a common name (Ken Smith). Many other attributes are possible, so you can send e-mail to someone whose exact e-mail address you do not know, provided you know enough other attributes (e.g., company and job title). Although X.400 names are considerably less convenient than DNS names, most e-mail systems have aliases (sometimes called nicknames) that allow users to enter or select a person's name and get the correct e-mail address. Consequently, even with X.400 addresses, it is usually not necessary to actually type in these strange strings.

Most e-mail systems support mailing lists, so that a user can send the same message to a list of people with a single command. If the mailing list is maintained locally, the user agent can just send a separate message to each intended recipient. However, if the list is maintained remotely, then messages will be expanded there. For example, if a group of bird watchers has a mailing list called `birders` installed on `meadowlark.arizona.edu`, then

any message sent to `birders@meadowlark.arizona.edu` will be routed to the University of Arizona and expanded there into individual messages to all the mailing list members, wherever in the world they may be. Users of this mailing list cannot tell that it is a mailing list. It could just as well be the personal mailbox of Prof. Gabriel O. Birders.

Reading E-mail

Typically, when a user agent is started up, it looks at the user's mailbox for incoming e-mail before displaying anything on the screen. Then it may announce the number of messages in the mailbox or display a one-line summary of each one and wait for a command.

As an example of how a user agent works, let us take a look at a typical mail scenario. After starting up the user agent, the user asks for a summary of his e-mail. A display like that of figure then appears on the screen. Each line refers to one message. In this example, the mailbox contains eight messages.

Figure: An example display of the contents of a mailbox

#	Flags	Bytes	Sender	Subject
1	K	1030	asw	Changes to MINIX
2	KA	6348	trudy	Not all Trudys are nasty
3	K F	4519	Amy N. Wong	Request for information
4		1236	bal	Bioinformatics
5		104110	kaashoek	Material on peer-to-peer
6		1223	Frank	Re: Will you review a grant proposal
7		3110	guido	Our paper has been accepted
8		1204	dmr	Re: My student's visit

Each line of the display contains several fields extracted from the envelope or header of the corresponding message. In a simple e-mail system, the choice of fields displayed is built into the program. In a more sophisticated system, the user can specify which fields are to be displayed by providing a user profile, a file describing the display format. In this basic example, the first field is the message number. The second field, Flags, can contain a K, meaning that the message is not new but was read previously and kept in the mailbox; an A, meaning that the message has already been answered; and/or an F, meaning that the message has been forwarded to someone else. Other flags are also possible.

The third field tells how long the message is, and the fourth one tells who sent the message. Since this field is simply extracted from the message, this field may contain first names, full names, initials, login names, or whatever else the sender chooses to put there. Finally, the Subject field gives a brief summary of what the message is about.

People who fail to include a Subject field often discover that responses to their e-mail tend not to get the highest priority.

After the headers have been displayed, the user can perform any of several actions, such as displaying a message, deleting a message, and so on. The older systems were text based and typically used one-character commands for performing these tasks, such as T (type message), A (answer message), D (delete message), and F (forward message). An argument specified the message in question. More recent systems use graphical interfaces. Usually, the user selects a message with the mouse and then clicks on an icon to type, answer, delete, or forward it.

E-mail has come a long way from the days when it was just file transfer. Sophisticated user agents make managing a large volume of e-mail possible. For people who receive and send thousands of messages a year, such tools are invaluable.

2.3 Message Transfer

The message transfer system is concerned with relaying messages from the originator to the recipient. The simplest way to do this is to establish a transport connection from the source machine to the destination machine and then just transfer the message. After examining how this is normally done, we will examine some situations in which this does not work and what can be done about them.

SMTP— The Simple Mail Transfer Protocol

Within the Internet, e-mail is delivered by having the source machine establish a TCP connection to port 25 of the destination machine. Listening to this port is an e-mail daemon that speaks SMTP (Simple Mail Transfer Protocol). This daemon accepts incoming connections and copies messages from them into the appropriate mailboxes. If a message cannot be delivered, an error report containing the first part of the undeliverable message is returned to the sender.

SMTP is a simple ASCII protocol. After establishing the TCP connection to port 25, the sending machine, operating as the client, waits for the receiving machine, operating as the server, to talk first. The server starts by sending a line of text giving its identity and telling whether it is prepared to receive mail. If it is not, the client releases the connection and tries again later.

If the server is willing to accept e-mail, the client announces whom the e-mail is coming from and whom it is going to. If such a recipient exists at the destination, the server gives the client the go-ahead to send the message. Then the client sends the message and the server acknowledges it. No checksums are needed because TCP provides a reliable byte stream. If there is more e-mail, that is now sent. When all the e-mail has been exchanged in both directions, the connection is released. A sample dialog for sending the message including the numerical codes used by SMTP, is shown in figure. The lines sent by the client are marked C:. Those sent by the server are marked S:.

Figure: Transferring a message from elinor@abcd.com to carolyn@xyz.com

S: 220 xyz.com SMTP service ready
 C: HELO abcd.com
 S: 250 xyz.com says hello to abcd.com
 C: MAIL FROM: <elinor@abcd.com>
 S: 250 sender ok
 C: RCPT TO: <carolyn@xyz.com>
 S: 250 recipient ok
 C: DATA
 S: 354 Send mail; end with "." on a line by itself
 C: From: elinor@abcd.com
 C: To: carolyn@xyz.com
 C: MIME-Version: 1.0
 C: Message-Id: <0704760941.AA00747@abcd.com>
 C: Content-Type: multipart/alternative; boundary=qwertyuiopasdfghjklzxcvbnm
 C: Subject: Earth orbits sun integral number of times
 C:
 C: This is the preamble. The user agent ignores it. Have a nice day.
 C:
 C: --qwertyuiopasdfghjklzxcvbnm
 C: Content-Type: text/enriched
 C:
 C: Happy birthday to you
 C: Happy birthday to you
 C: Happy birthday dear <bold> Carolyn </bold>
 C: Happy birthday to you
 C:
 C: --qwertyuiopasdfghjklzxcvbnm
 C: Content-Type: message/external-body;
 C: access-type="anon-ftp";
 C: site="bicycle.abcd.com";
 C: directory="pub";
 C: name="birthday.snd"
 C:
 C: content-type: audio/basic
 C: content-transfer-encoding: base64
 C: --qwertyuiopasdfghjklzxcvbnm
 C:
 S: 250 message accepted
 C: QUIT
 S: 221 xyz.com closing connection

A few comments about the figure may be helpful. The first command from the client is indeed HELO. Of the various four-character abbreviations for HELLO, this one has numerous advantages over its biggest competitor. Why all the commands had to be four characters has been lost in the mists of time.

In the figure, the message is sent to only one recipient, so only one RCPT command is used. Such commands are allowed to send a single message to multiple receivers. Each one is individually acknowledged or rejected. Even if some recipients are rejected (because they do not exist at the destination), the message can be sent to the other ones.

Finally, although the syntax of the four-character commands from the client is rigidly specified, the syntax of the replies is less rigid. Only the numerical code really counts. Each implementation can put whatever string it wants after the code.

To get a better feel for how SMTP and some of the other protocols described in this chapter work, try them out. In all cases, first go to a machine connected to the Internet. On a UNIX system, in a shell, type

```
telnet mail.isp.com 25
```

substituting the DNS name of your ISP's mail server for mail.isp.com. On a Windows system, click on Start, then Run, and type the command in the dialog box. This command will establish a telnet (i.e., TCP) connection to port 25 on that machine. Port 25 is the SMTP port. You will probably get a response something like this:

```
Trying 192.30.200.66...
Connected to mail.isp.com
Escape character is '^]'.
220 mail.isp.com Smail #74 ready at Thu, 25 Sept 2002 13:26 +0200
```

The first three lines are from telnet telling you what it is doing. The last line is from the SMTP server on the remote machine announcing its willingness to talk to you and accept e-mail. To find out what commands it accepts, type

```
HELP
```

From this point on, a command sequence such as the one in figure is possible, starting with the client's HELO command.

It is worth noting that the use of lines of ASCII text for commands is not an accident. Most Internet protocols work this way. Using ASCII text makes the protocols easy to test and debug. They can be tested by sending commands manually, as we saw above, and dumps of the messages are easy to read.

Even though the SMTP protocol is completely well defined, a few problems can still arise. One problem relates to message length. Some older implementations cannot handle messages exceeding 64 KB. Another problem relates to timeouts. If the client and server have different timeouts, one of them may give up while the other is still busy, unexpectedly terminating the connection. Finally, in rare situations, infinite mail storms can be triggered. For example, if host 1 holds mailing list A and host 2 holds mailing list B and each list contains an entry for the other one, then a message sent to either list could generate a never-ending amount of e-mail traffic unless somebody checks for it.

To get around some of these problems, extended SMTP (ESMTP) has been defined. Clients wanting to use it should send an EHLO message instead of HELO initially. If this is rejected, then the server is a regular SMTP server, and the client should proceed in the usual way. If the EHLO is accepted, then new commands and parameters are allowed.

2.4 File Transfer Protocol (FTP)

FTP is used to copy files between computers over a TCP/IP network. FTP uses the transport services of TCP to provide a reliable, connection-oriented service. It is commonly used when downloading files from the Internet. FTP requires a valid user name and password to access a private FTP server. Anonymous FTP is used to access public FTP servers. With Anonymous FTP, the username is entered as "anonymous" and the password is the user's email address. The computer requesting a file from an FTP server must have an FTP client application installed and the server must have an FTP server application installed. Since FTP relies on the presence of TCP/IP on both the client and the server computers, FTP will work no matter what operating system is on either machine.

2.5 Telnet

The Telnet protocol supports remote log-in to a host from another host on the network. Telnet emulates a dumb terminal, much like the old computer terminals connected to mainframe computers. Dumb terminals are computers that have no processing capabilities. They rely on the processing capabilities of the computers to which they are connected. Telnet allows the emulated terminal to make a direct connection to the remote host computer. To make a telnet connection the user must know the domain name or IP address of the remote computer, the login name, and, if required, the password. Once a connection is established, any command typed on the client's keyboard is sent to the remote computer for execution. Telnet is used in a variety of ways including connecting to a device to configure the device, to control a remote computer, or to manage a remote web server.

2.6 Leased Line

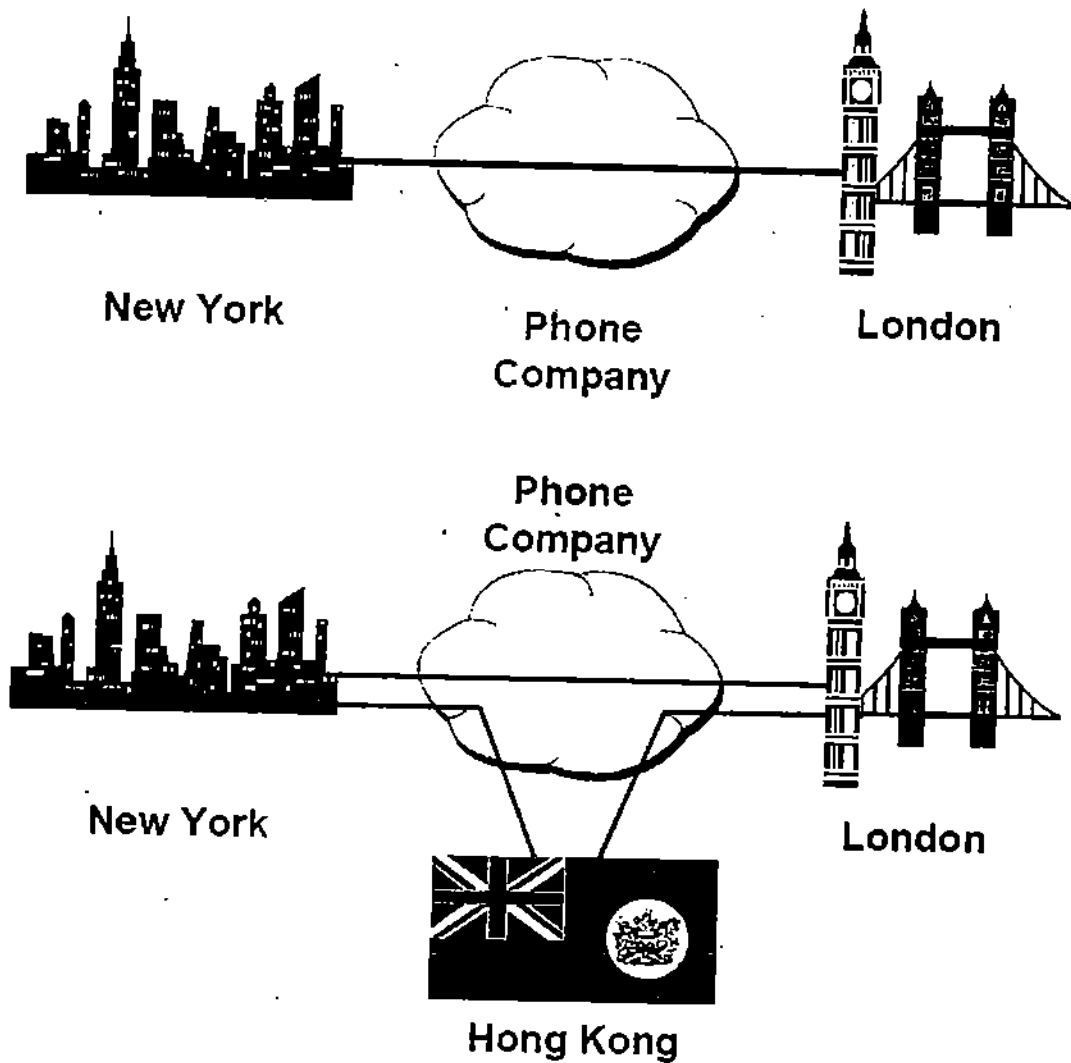
Leased Line (also called Dedicated Line or Private Line)~ A non-switched connection for a WAN. Only the company that leases it uses the line and the connection is always open.

When you make a regular voice phone call it can take as long as 30 seconds to a minute from the time you start dialing to when the phone at the other end starts ringing. Digital phone lines make connections much faster but there is still a wait while the circuit is established. All that waiting makes a dial-up connection a very slow way to transmit data.

A company that needs to share data constantly wouldn't want to wait for a circuit each time it must transmit data. For example, a stockbroker in London needs to see the prices of stocks on the New York Stock Exchange. S/he can't wait to dial the phone and make the connection; in that time the stock market may have changed.

Instead of using a dial-up connection, the broker might use a leased line (also called a dedicated line). A leased line is always connected, much like the way computers on a LAN are always connected. For each leased line, the bank pays the phone company for a circuit that stays open all of the time, like an endless phone call. This is expensive. Phone companies offer fractional T1, full T1, T3, and faster leased lines. The faster the speed and the greater the distance, the more a leased line will cost.

Point-to-Point Connections



A leased line connects through the phone company to one other point on the WAN. In the example, the leased line would connect the London office to New York. Like circuit switching, leased lines are point-to-point connections. If the bank also needed to connect an office in Hong Kong to New York, it would have to lease an additional line. And if it needed to connect London and Hong Kong, it would need a third leased line.

Some banks use leased lines for their automatic teller machines. Because the connection is always there, you don't need to wait for the ATM to call the bank to check your card number, password, and account balance. This is why you can get money so quickly. Security is another reason why banks use leased lines. You will learn more about security in the next lesson, WAN Considerations.

2.7 Summary

This unit introduces the Internet services ranging from e-mail, news, and remote login to file transfer. It also provides the history of Internet services. It gives an outlook to Internet service providers (ISPs). Then e-mail service is detailed by covering user agent, sending and reading e-mail, and message transfer. Finally, file transfer protocol (FTP), Telnet, and Leased Line is explained briefly.

2.8 Review Questions

1. Describe different Internet service.
2. What do you understand by e-mail?
3. Describe how an e-mail is send? Also describe how it is read.
4. What is message transfer? Explain SMTP.
5. What do you mean by file transfer protocol (FTP)?
6. Explain Telnet.
7. Describe leased lines. Draw a point-to-point connection.

Unit III – ISDN & Bridge-Routers

What You Will Learn (Objectives)

After completing this unit you will be able to:

- Define ISDN
- Know history and advantages of ISDN
- Describe ISDN Interfaces and protocols
- Explain Bridge-Routers

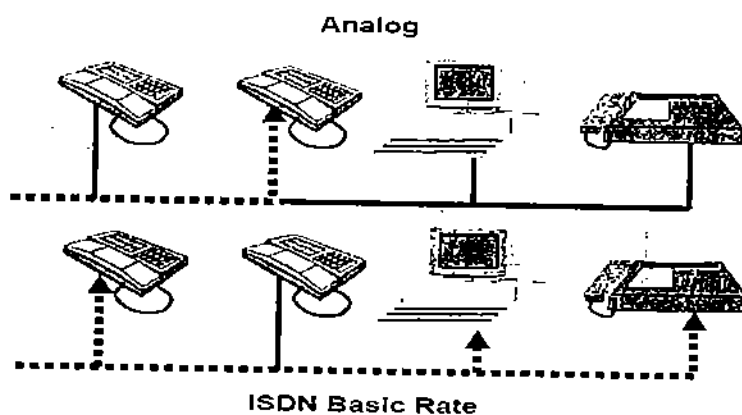
3.1 Introduction

Integrated Service Digital Network (ISDN) is digital telephone lines that can be used for circuit switched or packet switched data transmission. Integrated Services Digital Network (ISDN) is more than a kind of telephone line. It's a whole set of technologies including special devices such as telephones and switches, and standards that tell how data is to be transmitted. It has been around for a long time but hasn't become very popular in the United States. Different phone companies offered different versions and it was too confusing for most people. That situation is changing somewhat now, but it may be too late because newer, faster technologies are now available.

ISDN has certain advantages over T-carrier. T-carriers are frequently used by companies with lots of data to transmit, but even a fractional T1 is probably too expensive for an individual user who just wants to surf the Internet faster. Many people who work at home and many small businesses use *ISDN Basic Rate Interface* lines to connect to the Internet.

Like a T-carrier, ISDN-BRI uses the same twisted pair copper wire used for standard analog telephone lines and ISDN requires a device like a modem to convert the data from the digital format that comes from the computer to the digital format used on the ISDN lines. This device is called a *terminal adapter*.

Three Simultaneous Calls with ISDN. Only One with Analog



ISDN BRI divides the line into two 64 Kbps channels and a 16 Kbps channel that can be used separately for three different "conversations." ISDN lines can be used for regular phone conversations on the two 64 Kbps channels, as well as data transmission on any of the channels. The 64 Kbps channels can be combined to transmit data at 128 Kbps. This is an important advantage over T-carrier. ISDN channels can be combined and separated when needed. A graphic designer working from home with an ISDN line might use one channel for a telephone call to his/her office while the other channel is receiving a fax. Then s/he might combine the channels to access the Internet at 128 Kbps. In contrast, channels on a T-carrier are assigned to be used for voice calls or data transmission when the line is hooked up and can't be easily changed.

A faster form of ISDN called ISDN Primary Rate Interface (PRI) contains 23 64-Kbps and one additional 64 Kbps channel and can transmit data at up to 1.52 Mbps. This is suitable for larger business LANs with more data to transmit. It's no coincidence that T1 lines and ISDN-PRI lines have the same number of channels. ISDN PRI is designed to run over a T1.

The newest form of ISDN is Broadband ISDN (B-ISDN) that can handle data rates of 155-622 Mbps. B-ISDN has over 1000 times the bandwidth of regular ISDN, enough to carry cable television, interactive videoconferences, and other high-bandwidth data.

3.2 ISDN Definitions

ISDN, which stands for **I**ntegrated **S**ervices **D**igital **N**etwork, is a system of digital phone connections which has been available for over a decade. This system allows voice and data to be transmitted simultaneously across the world using end-to-end digital connectivity.

With ISDN, voice and data are carried by bearer channels (**B channels**) occupying a bandwidth of 64 kb/s (bits per second). Some switches limit B channels to a capacity of 56 kb/s. A data channel (**D channel**) handles signaling at 16 kb/s or 64 kb/s, depending on the service type. Note that, in ISDN terminology, "**k**" means 1000 (10^3), not 1024 (2^{10}) as in many computer applications (the designator "**K**" is sometimes used to represent this value); therefore, a 64 kb/s channel carries data at a rate of 64000 b/s. A new set of standard prefixes has recently been created to handle this. Under this scheme, "**k**" (kilo-) means 1000 (10^3), "**M**" (mega-) means 1000000 (10^6), and so on, and "**Ki**" (kibi-) means 1024 (2^{10}), "**Mi**" (mebi-) means 1048576 (2^{20}), and so on.

There are two basic types of ISDN service: **Basic Rate Interface (BRI)** and **Primary Rate Interface (PRI)**. BRI consists of two 64 kb/s B channels and one 16 kb/s D channel for a total of 144 kb/s. This basic service is intended to meet the needs of most individual users.

PRI is intended for users with greater capacity requirements. Typically the channel structure is 23 B channels plus one 64 kb/s D channel for a total of 1536 kb/s. In Europe, PRI consists of 30 B channels plus one 64 kb/s D channel for a total of 1984 kb/s. It is also possible to support multiple PRI lines with one 64 kb/s D channel using **Non-Facility Associated Signaling (NFAS)**.

H channels provide a way to aggregate B channels. They are implemented as:

- H0=384 kb/s (6 B channels)
- H10=1472 kb/s (23 B channels)
- H11=1536 kb/s (24 B channels)
- H12=1920 kb/s (30 B channels) - International (E1) only

To access BRI service, it is necessary to subscribe to an ISDN phone line. Customer must be within 18000 feet (about 3.4 miles or 5.5 km) of the telephone company central office for BRI service; beyond that, expensive repeater devices are required, or ISDN service may not be available at all. Customers will also need special equipment to communicate with the phone company switch and with other ISDN devices. These devices include **ISDN Terminal Adapters** (sometimes called, incorrectly, "ISDN Modems") and **ISDN Routers**.

3.3 ISDN History

The early phone network consisted of a pure analog system that connected telephone users directly by a mechanical interconnection of wires. This system was very inefficient, was very prone to breakdown and noise, and did not lend itself easily to long-distance connections. Beginning in the 1960s, the telephone system gradually began converting its internal connections to a packet-based, digital switching system. Today, nearly all voice switching in the U.S. is digital within the telephone network. Still, the final connection from the local central office to the customer equipment was, and still largely is, an analog Plain-Old Telephone Service (POTS) line.

A standards movement was started by the International Telephone and Telegraph Consultative Committee (CCITT), now known as the ITU. The ITU is an organization that coordinates and standardizes international telecommunications. Original recommendations of ISDN were in CCITT Recommendation I.120 (1984) which described some initial guidelines for implementing ISDN.

Local phone networks, especially the regional Bell operating companies, have long hailed the system, but they had been criticized for being slow to implement ISDN. One good reason for the delay is the fact that the two major switch manufacturers, Northern Telecom, and AT&T, selected different ways to implement the CCITT standards. These standards didn't always interoperate. This situation has been likened to that of earlier 19th century railroading. "People had different gauges, different tracks... nothing worked well."

In the early 1990s, an industry-wide effort began to establish a specific implementation for ISDN in the U.S. Members of the industry agreed to create the National ISDN 1 (NI-1) standard so that end users would not have to know the brand of switch they are connected to in order to buy equipment and software compatible with it. However, there were problems agreeing on this standard. In fact, many western states would not implement NI-1.

Ultimately, all the Regional Bell Operating Companies (RBOCs) did support NI-1. A more comprehensive standardization initiative, National ISDN 2 (NI-2), was later adopted. Some manufacturers of ISDN communications equipment, such as Motorola and US Robotics, worked with the RBOCs to develop configuration standards for their equipment. These kinds of actions, along with more competitive pricing, inexpensive ISDN connection equipment, and the desire for people to have relatively low-cost high-bandwidth Internet access have made ISDN more popular in recent years.

Most recently, ISDN service has largely been displaced by broadband internet service, such as xDSL and Cable Modem service. These services are faster, less expensive, and easier to set up and maintain than ISDN. Still, ISDN has its place, as backup to dedicated lines, and in locations where broadband service is not yet available.

3.4 Advantages of ISDN

Speed

The modem was a big breakthrough in computer communications. It allowed computers to communicate by converting their digital information into an analog signal to travel through the public phone network. There is an upper limit to the amount of information that an analog telephone line can hold. Currently, it is about 56 kb/s bidirectionally. Commonly available modems have a maximum speed of 56 kb/s, but are limited by the quality of the analog connection and routinely go about 45-50 kb/s. Some phone lines do not support 56 kb/s connections at all. There were currently 2 competing, incompatible 56 kb/s standards (X2 from U S Robotics (recently bought by 3Com), and K56flex from Rockwell/Lucent). This standards problem was resolved when the ITU released the V.90, and later V.92, standard for 56 kb/s modem communications.

ISDN allows multiple digital channels to be operated simultaneously through the same regular phone wiring used for analog lines. The change comes about when the telephone company's switches can support digital connections. Therefore, the same physical wiring can be used, but a digital signal, instead of an analog signal, is transmitted across the line. This scheme permits a much higher data transfer rate than analog lines. BRI ISDN, using a channel aggregation protocol such as BONDING or Multilink-PPP, supports an uncompressed data transfer speed of 128 kb/s, plus bandwidth for overhead and signaling. In addition, the latency, or the amount of time it takes for a communication to begin, on an ISDN line is typically about half that of an analog line. This improves response for interactive applications, such as games.

Multiple Devices

Previously, it was necessary to have a separate phone line for each device you wished to use simultaneously. For example, one line each was required for a telephone, fax, computer, bridge/router, and live video conference system. Transferring a file to someone while talking on the phone or seeing their live picture on a video screen would require several potentially expensive phone lines.

ISDN allows multiple devices to share a single line. It is possible to combine many different digital data sources and have the information routed to the proper destination. Since the line is digital, it is easier to keep the noise and interference out while combining these signals. ISDN technically refers to a specific set of digital services provided through a single, standard interface. Without ISDN, distinct interfaces are required instead.

Signaling

Instead of the phone company sending a ring voltage signal to ring the bell in your phone ("In-Band signal"), it sends a digital packet on a separate channel ("Out-of-Band signal"). The Out-of-Band signal does not disturb established connections, no bandwidth is taken from the data channels, and call setup time is very fast. For example, a V.90 or V.92 modem typically takes 30-60 seconds to establish a connection; an ISDN call setup usually takes less than 2 seconds.

The signaling also indicates who is calling, what type of call it is (data/voice), and what number was dialed. Available ISDN phone equipment is then capable of making intelligent decisions on how to direct the call.

3.5 Interfaces

In the U.S., the telephone company provides its BRI customers with a **U interface**. The U interface is a two-wire (single pair) interface from the phone switch, the same physical interface provided for POTS lines. It supports full-duplex data transfer over a single pair of wires, therefore only a single device can be connected to a U interface. This device is called a **Network Termination 1 (NT-1)**. The situation is different elsewhere in the world, where the phone company is allowed to supply the NT-1, and thereby the customer is given an **S/T interface**.

The NT-1 is a relatively simple device that converts the 2-wire U interface into the 4-wire **S/T interface**. The S/T interface supports multiple devices (up to 7 devices can be placed on the S/T bus) because, while it is still a full-duplex interface, there is now a pair of wires to receive data, and another to transmit data. Today, many devices have NT-1s built into their design. This has the advantage of making the devices less expensive and easier to install, but often reduces flexibility by preventing additional devices from being connected.

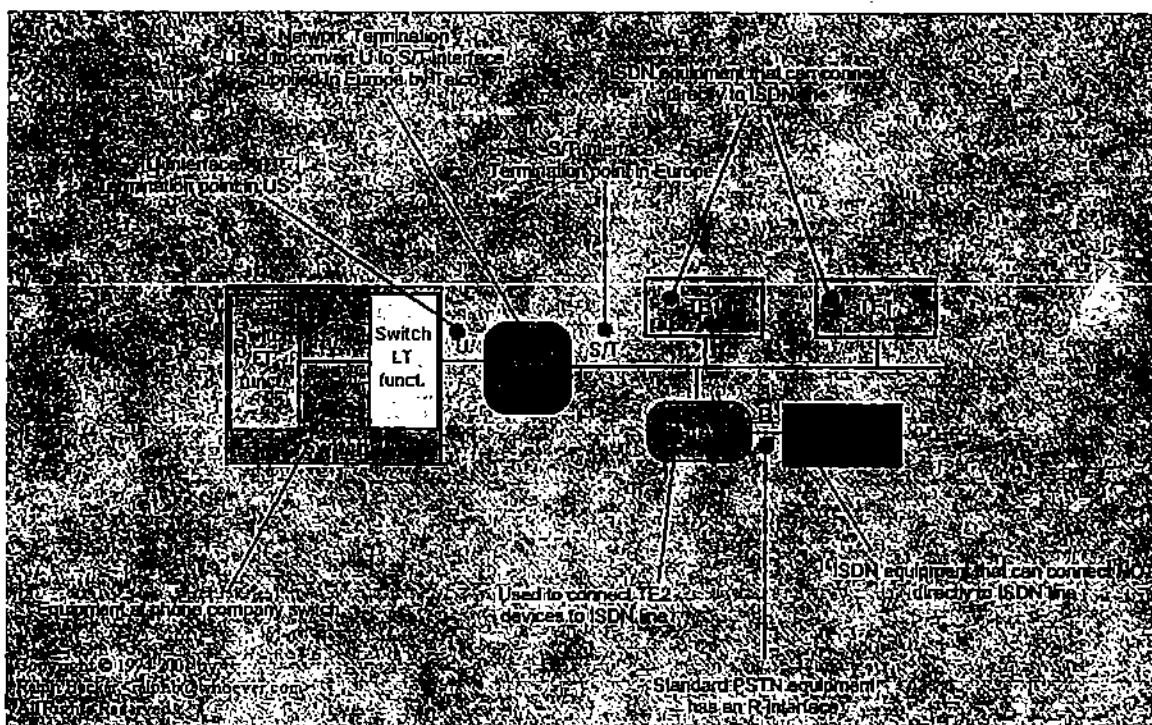
Technically, ISDN devices must go through an **Network Termination 2 (NT-2)** device, which converts the T interface into the S interface (Note: the S and T interfaces are electrically equivalent). Virtually all ISDN devices include an NT-2 in their design. The NT-2 communicates with terminal equipment, and handles the Layer 2 and 3 ISDN

protocols. Devices most commonly expect either a U interface connection (these have a built-in NT-1), or an S/T interface connection.

Devices that connect to the S/T (or S) interface include ISDN capable telephones and FAX machines, video teleconferencing equipment, bridge/routers, and terminal adapters. All devices that are designed for ISDN are designated **Terminal Equipment 1 (TE1)**. All other communication devices that are *not* ISDN capable, but have a POTS telephone interface (also called the **R interface**), including ordinary analog telephones, FAX machines, and modems, are designated **Terminal Equipment 2 (TE2)**. A **Terminal Adapters (TA)** connects a TE2 to an ISDN S/T bus.

Going one step in the opposite direction takes us inside the telephone switch. Remember that the U interface connects the switch to the customer premises equipment. This local loop connection is called *Line Termination* (LT function). The connection to other switches within the phone network is called *Exchange Termination* (ET function). The LT function and the ET function communicate via the **V interface**.

This can get rather confusing. This diagram should be helpful:



3.6 Physical Layer Protocols

The ISDN Physical Layer is specified by the ITU I-series and G-series documents. The U interface provided by the telco for BRI is a 2-wire, 160 kb/s digital connection. Echo cancellation is used to reduce noise, and data encoding schemes (2B1Q in North America, 4B3T in Europe) permit this relatively high data rate over ordinary single-pair local loops.

2B1Q

2B1Q (2 Binary 1 Quaternary) is the most common signaling method on U interfaces. This protocol is defined in detail in 1988 ANSI spec T1.601. In summary, 2B1Q provides:

- Two bits per baud
- 80 kilobaud (baud = 1 modulation per second)
- Transfer rate of 160 kb/s

Bits	Quaternary Symbol	Voltage Level
00	-3	-2.5
01	-1	-0.833
10	+3	+2.5
11	+1	+0.833

This means that the input voltage level can be one of 4 distinct levels (note: 0 volts is not a valid voltage under this scheme). These levels are called **Quaternaries**. Each quaternary represents 2 data bits, since there are 4 possible ways to represent 2 bits, as in the table above.

Frame Format

Each U interface frame is 240 bits long. At the prescribed data rate of 160 kb/s, each frame is therefore 1.5 ms long. Each frame consists of:

- Frame overhead - 16 kb/s
- D-channel - 16 kb/s
- 2 B channels at 64 kb/s - 128 kb/s

Sync 18 bits	12 * (B₁ + B₂ + D) 216 bits	Maintenance 6 bits
------------------------	---	------------------------------

- The Sync field consists of 9 Quaternaries (2 bits each) in the pattern +3 +3 -3 -3 -3 +3 -3 +3 -3.
- (B₁ + B₂ + D) is 18 bits of data consisting of 8 bits from the first B channel, 8 bits from the second B channel, and 2 bits of D channel data.

The Maintenance field contains CRC information, block error detection flags, and "embedded operator commands" used for loopback testing without disrupting user data.

Data is transmitted in a **superframe** consisting of 8 240-bit frames for a total of 1920 bits (240 octets). The sync field of the first frame in the superframe is inverted (i.e. -3 -3 +3 +3 +3 -3 +3 -3 +3).

3.7 Link Layer Protocols

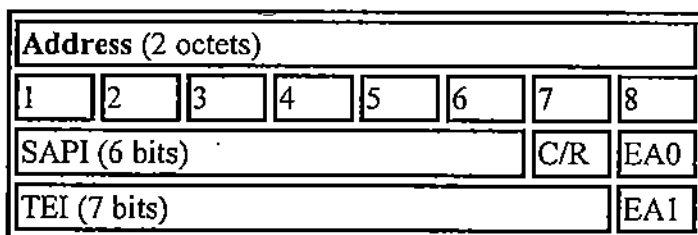
The ISDN Data Link Layer is specified by the ITU Q-series documents Q.920 through Q.923. All of the signaling on the D channel is defined in the Q.921 spec.

LAP-D

Link Access Protocol - D channel (LAP-D) is the Layer 2 protocol used. This is almost identical to the X.25 LAP-B protocol. Here is the structure of a LAP-D frame:



Flag (1 octet) - This is always 7E₁₆ (0111 1110₂)



SAPI (Service access point identifier), 6-bits (see below)

C/R (Command/Response) bit indicates if the frame is a command or a response

EA0 (Address Extension) bit indicates whether this is the final octet of the address or not

TEI (Terminal Endpoint Identifier) 7-bit device identifier (see below)

EA1 (Address Extension) bit, same as EA0

Control (2 octets) - The frame level control field indicates the frame type (Information, Supervisory, or Unnumbered) and sequence numbers (N(r) and N(s)) as required.

Information - Layer 3 protocol information and User data

CRC (2 octets) - Cyclic Redundancy Check is a low-level test for bit errors on the user data.

Flag (1 octet) - This is always 7E₁₆ (0111 1110₂)

SAPIs

The Service Access Point Identifier (SAPI) is a 6-bit field that identifies the point where Layer 2 provides a service to Layer 3. See the following table:

SAPI	Description
0	Call control procedures
1	Packet Mode using Q.931 call procedures
16	Packet Mode communications procedures
32-47	Reserved for national use
63	Management Procedures
Others	Reserved for Future Use

TEIs

Terminal Endpoint Identifiers (TEIs) are unique IDs given to each device (TE) on an ISDN S/T bus. This identifier can be dynamic; the value may be assigned statically when the TE is installed, or dynamically when activated.

TEI	Description
0-63	Fixed TEI assignments
64-126	Dynamic TEI assignment (assigned by the switch)
127	Broadcast to all devices

Establishing the Link Layer

The Layer 2 establishment process is very similar to the X.25 LAP-B setup, if you are familiar with it.

1. The TE (Terminal Endpoint) and the Network initially exchange Receive Ready (RR) frames, listening for someone to initiate a connection
2. The TE sends an Unnumbered Information (UI) frame with a SAPI of 63 (management procedure, query network) and TEI of 127 (broadcast)
3. The Network assigns an available TEI (in the range 64-126)
4. The TE sends a Set Asynchronous Balanced Mode (SABME) frame with a SAPI of 0 (call control, used to initiate a SETUP) and a TEI of the value assigned by the network
5. The network responds with an Unnumbered Acknowledgement (UA), SAPI=0, TEI=assigned.

At this point, the connection is ready for a Layer 3 setup.

3.8 Network Layer Protocols

The ISDN Network Layer is also specified by the ITU Q-series documents Q.930 through Q.939. Layer 3 is used for the establishment, maintenance, and termination of logical network connections between two devices.

SPIDs

Service Profile IDs (SPIDs) are used to identify what services and features the telco switch provides to the attached ISDN device. SPIDs are optional; when they are used, they are only accessed at device initialization time, before the call is set up. The format of the SPID is defined in a recommendation document, but it is only rarely followed. It is usually the 10-digit phone number of the ISDN line, plus a prefix and a suffix that are sometimes used to identify features on the line, but in reality it can be whatever the telco decides it should be. If an ISDN line requires a SPID, but it is not correctly supplied, then Layer 2 initialization will take place, but Layer 3 will not, and the device will not be able to place or accept calls. See ITU-spec Q.932 for details.

Information Field Structure

The Information Field is a variable length field that contains the Q.931 protocol data.

Information Field							
1	2	3	4	5	6	7	8
Protocol Discriminator							
0	0	0	0	Length of CRV			
Call Reference Value (1 or 2 octets)							
0	Message Type						
Mandatory				&	Optional		
Information Elements (variable)							

These are the fields in a Q.931 header: **Protocol Discriminator** (1 octet) - identifies the Layer 3 protocol. If this is a Q.931 header, this value is always 08₁₆. **Length** (1 octet) - indicates the length of the next field, the CRV. **Call Reference Value (CRV)** (1 or 2 octets) - used to uniquely identify each call on the user-network interface. This value is assigned at the beginning of a call, and this value becomes available for another call when the call is cleared. **Message Type** (1 octet) - identifies the message type (i.e., SETUP, CONNECT, etc.). This determines what additional information is required and allowed. **Mandatory and Optional Information Elements** (variable length) - are options that are set depending on the Message Type.

Layer 3 Call Setup

These are the steps that occurs when an ISDN call is established. In the following example, there are three points where messages are sent and received; 1) the Caller, 2) the ISDN Switch, and 3) the Receiver.

1. Caller sends a SETUP to the Switch.
2. If the SETUP is OK, the switch sends a CALL PROCEEDing to the Caller, and then a SETUP to the Receiver.
3. The Receiver gets the SETUP. If it is OK, then it rings the phone and sends an ALERTING message to the Switch.

4. The Switch forwards the ALERTING message to the Caller.
5. When the receiver answers the call, it sends a CONNECT message to the Switch
6. The Switch forwards the CONNECT message to the Caller.
7. The Caller sends a CONNECT ACKnowledge message to the Switch
8. The Switch forwards the CONNECT ACK message to the Receiver.
9. Done. The connection is now up.

3.9 Bridge-Routers

The early designs for bridges were intended for use between local area networks (LANs) that use identical protocols for the physical and medium access layers (e.g., all conforming to IEEE 802.3 or all conforming to FDDI). Because the devices all use the same protocols, the amount of processing required at the bridge is minimal. In recent years, bridges that operate between LANs with different MAC protocols have been developed.

Because the bridge is used in a situation in which all of the LANs have the same characteristics, the reader may ask why one does not simply use one large LAN. Depending on circumstance, there are several reasons for the use of multiple LANs connected by bridges:

- Reliability
- Performance
- Security
- Geography

3.9.1 Bridge Function

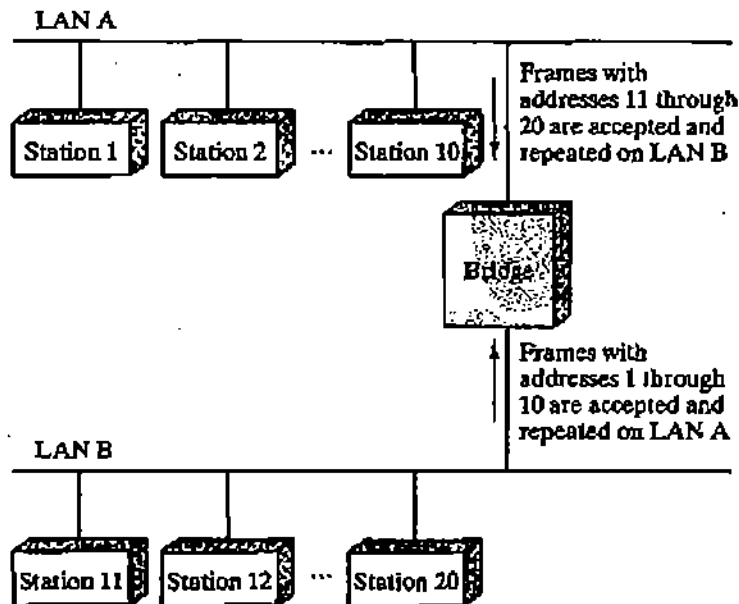
Figure (a) illustrates the operation of a bridge between two LANs, A and B. The bridge performs the following functions:

- * Reads all frames transmitted on A, and accepts those addressed to stations on B.
- * Using the medium access control protocol for B, retransmits the frames onto B.
- * Does the same for B-to-A traffic.

Several design aspects of a bridge are worth highlighting:

1. The bridge makes no modification to the content or format of the frames it receives, nor does it encapsulate them with an additional header. Each frame to be transferred is simply copied from one LAN and repeated with exactly the same bit pattern as the other LAN. Because the two LANs use the same LAN protocols, it is permissible to do this.

Figure (a)



2. The bridge should contain enough buffer space to meet peak demands. Over a short period of time, frames may arrive faster than they can be retransmitted.

3. The bridge must contain addressing and routing intelligence. At a minimum, the bridge must know which addresses are on each network in order to know which frames to pass. Further, there may be more than two LANs interconnected by a number of bridges. In that case, a frame may have to be routed through several bridges in its journey from source to destination.

4. A bridge may connect more than two LANs.

The bridge provides an extension to the LAN that requires no modification to the communications software in the stations attached to the LANs. It appears to all stations on the two (or more) LANs that there is a single LAN on which each station has a unique address. The station uses that unique address and need not explicitly discriminate between stations on the same LAN and stations on other LANs; the bridge takes care of that.

The description above has applied to the simplest sort of bridge. More sophisticated bridges can be used in more complex collections of LANs. These constructions would include additional functions, such as,

A. Each bridge can maintain status information on other bridges, together with the cost and number of bridge-to-bridge hops required to reach each network. This information may be updated by periodic exchanges of information among bridges; this allows the bridges to perform a dynamic routing function.

B. A control mechanism can manage frame buffers in each bridge to overcome congestion. Under saturation conditions, the bridge can give precedence to en route

packets over new packets just entering the internet from an attached LAN, thus preserving the investment in line bandwidth and processing time already made in the en route frame.

3.9.2 Routing with Bridges

In the configuration of Figure (a), the bridge makes the decision to relay a frame on the basis of destination MAC address. In a more complex configuration, the bridge must also make a routing decision. Consider the configuration of Figure (b). Suppose that station 1 transmits a frame on LAN A intended for station 5. The frame will be read by both bridge 101 and bridge 102. For each bridge, the addressed station is not on a LAN to which the bridge is attached. Therefore, each bridge must make a decision of whether or not to retransmit the frame on its other LAN, in order to move it closer to its intended destination. In this case, bridge 101 should repeat the frame on LAN B, whereas bridge 102 should refrain from retransmitting the frame. Once the frame has been transmitted on LAN B, it will be picked up by both bridges 103 and 104. Again, each must decide whether or not to forward the frame. In this case, bridge 104 should retransmit the frame on LAN E, where it will be received by the destination, station 5.

Thus, we see that, in the general case, the bridge must be equipped with a routing capability. When a bridge receives a frame, it must decide whether or not to forward it. If the bridge is attached to two or more networks, then it must decide whether or not to forward the frame and, if so, on which LAN the frame should be transmitted.

The routing decision may not always be a simple one. In Figure (c), bridge 107 is added to the previous configuration, directly linking LAN A and LAN E. Such an addition may be made to provide for higher overall internet availability. In this case, if Station 1 transmits a frame on LAN A intended for station 5 on LAN E, then either bridge 101 or bridge 107 could forward the frame. It would appear preferable for bridge 107 to forward the frame, as it will involve only one hop, whereas if the frame travels through bridge 101, it must suffer two hops. Another consideration is that there may be changes in the configuration. For example, bridge 107 may fail, in which case subsequent frames from station 1 to station 5 should go through bridge 101. We can say, then, that the routing capability must take into account the topology of the internet configuration and may need to be dynamically altered.

Figure (b) Configuration of Bridges & LANs

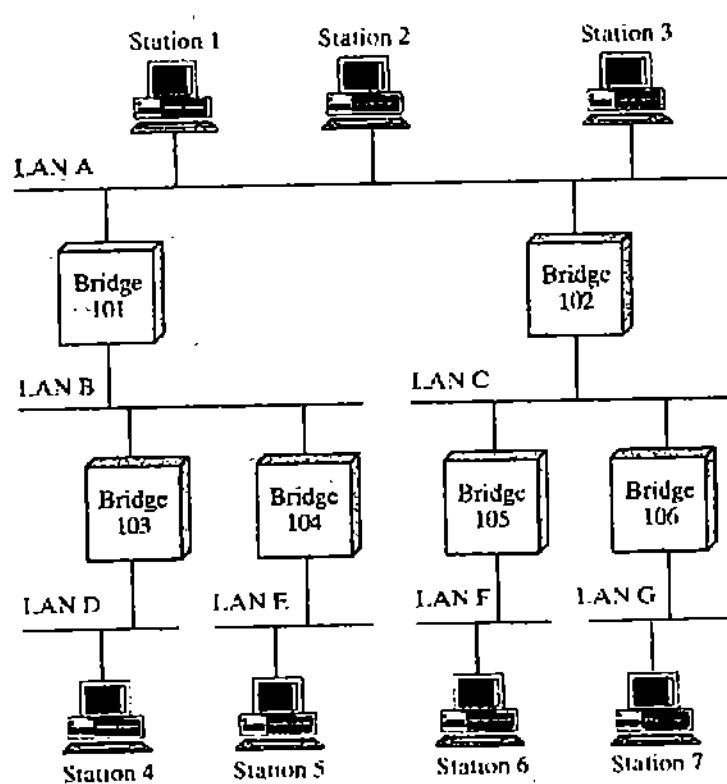
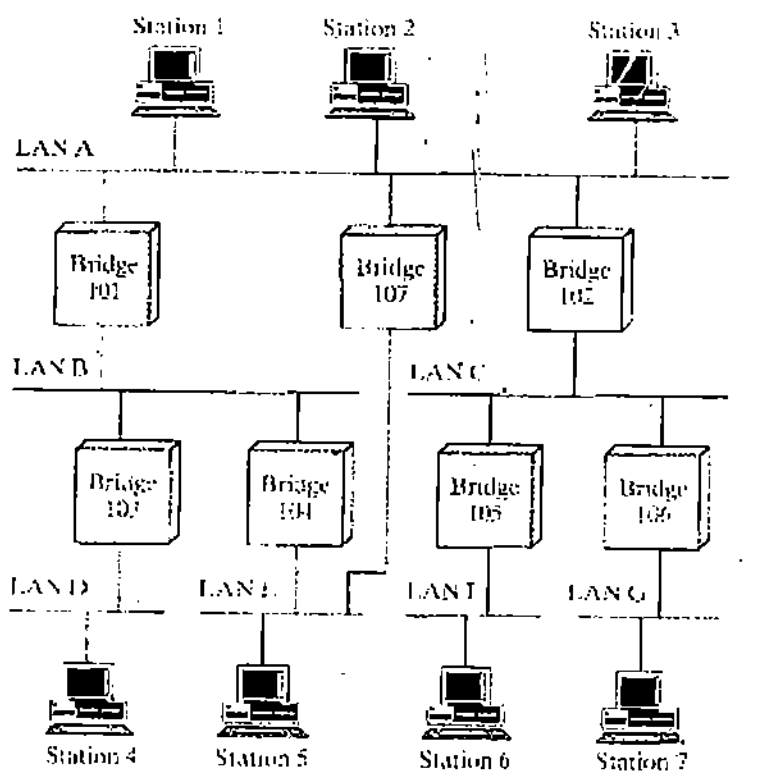


Figure (c) suggests that a bridge knows the identity of each station on each LAN. In a large configuration, such an arrangement is unwieldy. Furthermore, as stations are added to and dropped from LANs, all directories of station location must be updated. It would facilitate the development of a routing capability if all MAC-level addresses were in the form of a network part and a station part. For example, the IEEE 802.5 standard suggests that 16-bit MAC addresses consist of a 7-bit LAN number and an 8-bit station number, and that 48-bit addresses consist of a 14-bit LAN number and a 32-bit station number.¹ In the remainder of this discussion, we assume that all MAC addresses include a LAN number and that routing is based on the use of that portion of the address only.

A variety of routing strategies have been proposed and implemented in recent years. The simplest and most common strategy is *fixed routing*. This strategy is suitable for small LAN collections and for interconnections that are relatively stable. More recently, two groups within the IEEE 802 committee have developed specifications for routing strategies. The IEEE 802.1 group has issued a standard for routing based on the use of a *spanning tree* algorithm. The token ring committee, IEEE 802.5, has issued its own specification, referred to as *source routing*.

Figure (c) Configuration of Bridges & LANs, with alternate routes



3.10 Summary

This unit defines Integrated Services Digital Networks (ISDN) and provides history and advantages in terms of speed, multiple devices, and signaling. We learn ISDN interfaces and protocols (Physical Layer, Link Layer, and Network Layer). Then Bridge functioning is discussed with routing functionalities. Finally, Bridge-routers are elaborated.

3.11 Review Questions

1. What is ISDN? Define.
2. Describe the history of ISDN.
3. What are the advantages of ISDN in terms of speed, multiple devices, and signaling?
4. Describe different ISDN interfaces.
5. Explain physical layer ISDN protocols.
6. Explain link layer ISDN protocols.
7. Explain network layer ISDN protocols.
8. What are Bridge-Routers?
9. Describe the functioning of a Bridge.
10. Explain how routing is done with Bridges.

UNIT 4 – ISP Connectivity

What You Will Learn (Objectives)

After completing this unit you will be able to:

- Learn about Internet Service Providers (ISPs).
- Explain Cable Modem, DSL, SHDSL
- Know how ISP connect to the Internet
- Describe Broadband access and Leased Lines

4.1 Internet Service Provider (ISP)

An **Internet service provider** (abbr. **ISP**, also called **Internet access provider** or **IAP**) is a business or organization that provides consumers or businesses access to the Internet and related services. In the past, most ISPs were run by the phone companies. Now, ISPs can be started by just about any individual or group with sufficient money and expertise. In addition to Internet access via various technologies such as dial-up and DSL, they may provide a combination of services including Internet transit, domain name registration and hosting, web hosting, and collocation.

The internet started off as a closed network between government research laboratories and relevant parts of universities. It became popular and then universities and colleges started giving more of their member's access to it. As a result, commercial Internet Service Providers occurred to provide access for mainly those who missed their university accounts. In 1989 Brookline, Massachusetts-based **The World (internet service provider)** become the 1st commercial ISP. **Telerama** founded in 1991 claims to be the world's 3rd ISP.

4.2 ISP Connection Options

ISPs employ a range of technologies to enable consumers to connect to their network. For "home users", the most popular options include dial-up, DSL (typically ADSL), Broadband wireless access, Cable modem, and ISDN (typically BRI). For customers who have more demanding requirements, such as medium-to-large businesses, or other ISPs, DSL (often SHDSL or ADSL), Ethernet, Metro Ethernet, Gigabit Ethernet, Frame Relay, ISDN (BRI or PRI), ATM, satellite Internet access and SONET are more likely. With the increasing popularity of downloading music and online video and the general demand for faster page loads, higher bandwidth connections are becoming more popular.

4.2.1 Typical Home User Connection

1. Dial-up Internet Access

Dial-up access is a form of Internet access via telephone line. The client uses a modem connected to a computer and a telephone line to dial into an Internet service provider's (ISP) node to establish a modem-to-modem link, which is then routed to the Internet.

Availability: Dial-up requires no additional infrastructure on top of the telephone network. As telephone points are available throughout the world, dial-up remains useful to travelers. Dial-up

is usually the only choice available for most rural or remote areas where getting a broadband connection is impossible due to low population and demand. Sometimes dial-up access may also be an alternative to people who have limited budgets as it is offered for free by some, though broadband is now increasingly available at lower prices in countries such as the United States, Canada, Hong Kong, Australia and the United Kingdom due to market competition.

Dial-up requires time to establish a telephone connection (several seconds, depending on the location) and perform handshaking before data transfers can take place. In locales with telephone connection charges, each connection incurs an incremental cost. If calls are time-charged, the duration of the connection incurs costs.

Dial-up access is a transient connection, because either the user or the ISP terminates the connection. Internet service providers will often set a limit on connection durations to prevent hogging of access, and will disconnect the user — requiring reconnection and the costs and delays associated with it.

In recent years, the availability of dialup access numbers has been shrinking. Most notably in part due to consolidations within the primary U.S. modem network operators such as MCI, Qwest, Sprint, and Level3, who have refocused their attention to broadband Internet access services.

Performance: Modern dial-up modems typically have a maximum theoretical speed of 56 kbit/s (using the V.90 or V.92 protocol), although in most cases 40-50 kbit/s is the norm. Factors such as phone line noise as well as the quality of the modem itself play a large part in determining connection speeds. Some connections may be as low as 20 kbit/s in extremely noisy environments, such as in a hotel room where the phone line is shared with many, many extensions.

Dial-up connections usually have high latency that can be as high as 400 ms or even more, which can make online gaming or video conferencing difficult, if not impossible. First person shooter style games are the most sensitive to latency, making playing them impractical on dial-up, however some games such as *Star Wars: Galaxies*, *The Sims Online*, *Warcraft 3*, *Guild Wars*, and *Unreal Tournament* are capable of running on 56 K dial-up.

2. Digital Subscriber Line (DSL)

DSL or xDSL, is a family of technologies that provide digital data transmission over the wires of a local telephone network. DSL originally stood for **digital subscriber loop**, although in recent years, many have adopted **digital subscriber line** as a more marketing-friendly term for the most popular version of consumer-ready DSL, ADSL. DSL uses high frequency, regular telephone uses low frequency.

Typically, the download speed of consumer DSL services ranges from 512 kilobits per second (kbit/s) to 24,000 kbit/s, depending on DSL technology, line conditions and service level implemented. Typically, upload speed is lower than download speed for Asymmetric Digital

Subscriber Line (ADSL) and equal to download speed for Symmetric Digital Subscriber Line (SDSL).

Some variants of DSL connections, like ADSL and very high speed DSL (VDSL), typically work by dividing the frequencies used in a single phone line into two primary 'bands'. The ISP data is carried over the high frequency band (25 kHz and above) whereas the voice is carried over the lower frequency band (4 kHz and below). (See the ADSL article on how the 'high frequency band is sub-divided). The user typically installs a DSL filter on each phone. This filters out the high frequencies from the phone, so that the phone only sends or receives the lower frequencies (the human voice). The DSL modem and the normal telephone equipment can be used simultaneously on the line without interference from each other.

DSL Technologies

The line length limitations from telephone exchange to subscriber are more restrictive for higher data transmission rates. Technologies such as VDSL provide very high speed, short-range links as a method of delivering "triple play" services (typically implemented in fiber to the curb network architectures). Technologies like GDSL can further increase the data rate of DSL.

Example DSL technologies (sometimes called **xDSL**) include:

- High Data Rate Digital Subscriber Line (**HDSL**), also covered in this article
- Symmetric Digital Subscriber Line (SDSL), a standardized version of HDSL
- Asymmetric Digital Subscriber Line (ADSL), a version of DSL with a slower upload speed
- ISDN Digital Subscriber Line (IDSL)
- Rate-Adaptive Digital Subscriber Line (RADSL)
- Very High Speed Digital Subscriber Line (VDSL)
- Very High Speed Digital Subscriber Line 2 (VDSL2), an improved version of VDSL
- Symmetric High-speed Digital Subscriber Line (G.SHDSL), a standardized replacement for early proprietary SDSL by the International Telecommunication Union Telecommunication Standardization Sector
- Powerline Digital Subscriber Line (PDSL), a high speed power line communications solution which modulates high speed data onto existing electricity distribution infrastructure
- UDSL
- Etherloop Ethernet Local Loop
- GDSL Gigabit DSL, based on binder MIMO technologies.

3. Broadband Wireless Access

According to the 802.16-2004 standard, broadband means 'having instantaneous bandwidth greater than around 1 MHz and supporting data rates greater than about 1.5 Mbit/s. This means

that *Wireless Broadband* features speeds roughly equivalent to wired broadband access, such as that of ADSL or a cable modem.

Technology & Speeds

Few WISP's provide download speeds of over 100 Mbit/s; most broadband wireless access services are estimated to have a range of 50 km (30 miles) from a tower. Technologies used include LMDS and MMDS, as well as heavy use of the ISM bands and one particular access technology is being standardized by IEEE 802.16, also known as WiMAX. WiMAX is highly popular in Europe but has not met full acceptance in the United States because cost of deployment does not meet return on investment figures. In 2005 the Federal Communications Commission adopted a Report and Order that revised the FCC's rules to open the 3650 MHz band for terrestrial wireless broadband operations. On November 14, 2007 the Commission released Public Notice DA 07-4605 in which the Wireless Telecommunications Bureau announced the start date for licensing and registration process for the 3650-3700 MHz band.

Initially, Wireless Internet Service Providers (WISPs) were only found in rural areas not covered by cable or DSL. These early WISPs would employ a high-capacity T-carrier, such as a T1 or DS3 connection, and then broadcast the signal from a high elevation, such as at the top of a water tower. To receive this type of Internet connection, consumers mount a small dish to the roof of their home or office and point it to the transmitter. Line of sight was usually necessary for this type of technology, but technologies by Motorola have not adhered to this general rule.

4. Cable MODEM

A **cable modem** is a type of modem that provides access to a data signal sent over the cable television infrastructure. Cable modems are primarily used to deliver broadband Internet access in the form of cable internet, taking advantage of unused bandwidth on a cable television network. They are commonly found in Australia, New Zealand, Canada, Europe, the United Kingdom, Costa Rica, and the United States. In USA alone there were 22.5 million cable modem users during the first quarter of 2005, up from 17.4 million in the first quarter of 2004.

In network topology, a cable modem is a network bridge that conforms to IEEE 802.1D for Ethernet networking (with some modifications). The cable modem bridges Ethernet frames between a customer LAN and the coax cable network.

With respect to the OSI model, a cable modem is a data link layer (or layer 2) forwarder, rather than simply a modem.

A cable modem does support functionalities at other layers. In physical layer (or layer 1), the cable modem supports the Ethernet PHY on its LAN interface, and a DOCSIS defined cable-specific PHY on its HFC cable interface. It is to this cable-specific PHY that the name *cable modem* refers. In the network layer (or layer 3), the cable modem is a IP host in that it has its own IP address used by the network operator to manage and troubleshoot the device. In the transport layer (or layer 4) the cable modem supports UDP in association with its own IP

address, and it supports filtering based on TCP and UDP port numbers to, for example, block forwarding of NetBIOS traffic out of the customer's LAN. In the application layer (layer 5 or layer 7), the cable modem supports certain protocols that are used for management and maintenance, notably DHCP, SNMP, and TFTP.

Some cable modem devices may incorporate a router along with the cable modem functionality, to provide the LAN with its own IP network addressing. From a data forwarding and network topology perspective, this router functionality is typically kept distinct from the cable modem functionality (at least logically) even though the two may share a single enclosure and appear as one unit. So, the cable modem function will have its own IP address and MAC address as will the router.

4.2.2 Typical Business Connections

1. **DSL**
2. **SHDSL**

Single-Pair high-speed digital subscriber line (SHDSL) is a telecommunications technology for Digital Subscriber Line (DSL) subscriber lines. It describes a transmission method for signals on copper pair lines, being mostly used in access networks to connect subscribers to Telephone exchanges or POP Access Points.

G.SHDSL was standardized in February 2001 internationally by ITU-T with recommendation G.991.2.

G.SHDSL features symmetrical data rates from 192 kbit/s to 2,304 kbit/s of payload in 64 kbit/s increments for one pair and 384 kbit/s to 4,608 kbit/s in 128 kbit/s increments for two pair applications. The reach varies according to the loop rate and noise conditions (more noise or higher rate means decreased reach) and may be up to 3,000 meters. The two pair feature may alternatively be used for increased reach applications by keeping the data rate low (halving the data rate per pair will provide similar speeds to single pair lines while increasing the error/noise tolerance).

The payload may be either 'clear channel' (unstructured), T1 or E1 (full rate or fractional), n x ISDN Basic Rate Access (BRA), Asynchronous Transfer Mode (ATM) or 'dual bearer' mode (i.e. a mixture of two separate streams (e.g. T1 and 'packet based') sharing the payload bandwidth of the G.shdsl loop).

In Europe, a variant of G.SHDSL was standardized by ETSI using the name 'SDSL'. This ETSI variant is not compatible with the ITU-T G.SHDSL standardized regional variant for Europe and must not be confused with the usage of the term 'SDSL' in North America.

The latest standardization efforts (G.SHDSL.bis) tend to allow for flexibly changing the amount of bandwidth dedicated to each transport unit to provide 'dynamic rate repartitioning' of

bandwidth demands during the uptime of the interface and optionally provides for 'extended data rates' by using a different modulation method (32-TCPAM instead of 16-TCPAM, where TCPAM is Trellis-Coded Pulse Amplitude Modulation). Also, a new payload type is introduced: packet based, e.g. to allow for Ethernet-frames to be transported natively. (Currently, they may only be framed in ATM or T1/E1/...). G.SHDSL.bis can deliver a minimum of 2 Mbit/s and a maximum of 5.69 Mbit/s over distances of up to 2.7 km (9 Kft).

3. Ethernet

Ethernet is a family of frame-based computer networking technologies for local area networks (LANs). The name comes from the physical concept of the ether. It defines a number of wiring and signaling standards for the physical layer, through means of network access at the Media Access Control (MAC)/Data Link Layer, and a common addressing format.

Ethernet is standardized as IEEE 802.3. The combination of the twisted pair versions of Ethernet for connecting end systems to the network, along with the fiber optic versions for site backbones, is the most widespread wired LAN technology. It has been in use from the 1990s to the present, largely replacing competing LAN standards such as token ring, FDDI, and ARCNET. In recent years, Wi-Fi, the wireless LAN standardized by IEEE 802.11, is prevalent in home and small office networks and augmenting Ethernet in larger installations.

4.3 How ISP Connect to the Internet

Just as their customers pay them for Internet access, ISPs themselves pay upstream ISPs for Internet access. In the simplest case, a single connection is established to an upstream ISP using one of the technologies described above, and the ISP uses this connection to send or receive any data to or from parts of the Internet beyond its own network; in turn, the upstream ISP uses its own upstream connection, or connections to its other customers (usually other ISPs) to allow the data to travel from source to destination.

In reality, the situation is often more complicated. For example, ISPs with more than one point of presence (PoP) may have separate connections to an upstream ISP at multiple PoPs, or they may be customers of multiple upstream ISPs and have connections to each one at one or more of their PoPs. ISPs may engage in peering, where multiple ISPs interconnect with one another at a peering point or Internet exchange point (IX), allowing the routing of data between their networks, without charging one another for that data - data that would otherwise have passed through their upstream ISPs, incurring charges from the upstream ISP. ISPs that require no upstream and have only customers and/or peers are called Tier 1 ISPs, indicating their status as ISPs at the top of the Internet hierarchy. Routers, switches, Internet routing protocols, and the expertise of network administrators all have a role to play in ensuring that data follows the best available route and that ISPs can "see" one another on the Internet.

4.4 Virtual ISP

A Virtual ISP (vISP) purchases services from another ISP (sometimes called a wholesale ISP or similar within this context) that allow the vISP's customers to access the Internet via one or more point of presence (PoPs) that are owned and operated by the wholesale ISP. There are various models for the delivery of this type of service; for example, the wholesale ISP could provide network access to end users via its dial-up modem PoPs or DSLAMs installed in telephone exchanges, and route, switch, and/or tunnel the end user traffic to the vISP's network, whereupon they may route the traffic toward its destination. In another model, the vISP does not route any end user traffic, and needs only provide AAA (Authentication, Authorization and Accounting) functions, as well as any "value-add" services like email or web hosting. Any given ISP may use their own PoPs to deliver one service, and use a vISP model to deliver another service, or, use a combination to deliver a service in different areas. The service provided by a wholesale ISP in a vISP model is distinct from that of an upstream ISP, even though in some cases, they may both be one and the same company. The former provides connectivity from the end user's premises to the Internet or to the end user's ISP, the latter provides connectivity from the end user's ISP to all or parts of the rest of the Internet.

A vISP can also refer to a completely automated white label service offered to anyone at no cost or for a minimal set-up fee. The actual ISP providing the service generates revenue from the calls and may also share a percentage of that revenue with the owner of the vISP. All technical aspects are dealt with leaving the owner of vISP with the task of promoting the service. This sort of service is however declining due to the popularity of unmetered internet access also known as flatrate.

4.5 Related Services

1. Broadband Access

Broadband Internet access, often shortened to just "broadband", is high-speed Internet access—typically contrasted with dial-up access over a modem.

Dial-up modems are generally only capable of a maximum bitrate of 56 kbit/s (kilobits per second) and require the full use of a telephone line—whereas broadband technologies supply at least double this speed and generally without disrupting telephone use.

Although various minimum speeds have been used in definitions of broadband, ranging up from 64 kbit/s up to 1.0 Mbit/s, the OECD report^[1] is typical in counting only download speeds equal to or faster than 256 kbit/s as broadband, and the US FCC use 200 kbit/s in their definition.

Speeds are defined in terms of *maximum download* because several common consumer broadband technologies such as ADSL are "asymmetric"—supporting much slower maximum upload speeds than download.

"Broadband penetration" is now treated as a key economic indicator.

Overview

Broadband is often called **high-speed Internet**, because it usually has a high rate of data transmission. In general, any connection to the customer of 256 kbit/s (0.256 Mbit/s) or more is considered broadband Internet. The International Telecommunication Union Standardization Sector (ITU-T) recommendation I.113 has defined broadband as a transmission capacity that is faster than primary rate ISDN, at 1.5 to 2 Mbit/s. The FCC definition of broadband is 200 kbit/s (0.2 Mbit/s) in one direction, and advanced broadband is at least 200 kbit/s in both directions. The Organization for Economic Co-operation and Development (OECD) has defined broadband as 256 kbit/s in at least one direction and this bit rate is the most common baseline that is marketed as "broadband" around the world. There is no specific bit rate defined by the industry, however, "broadband" can mean lower-bit rate transmission methods. Some Internet Service Providers (ISPs) use this to their advantage in marketing lower-bit rate connections as broadband.

In practice, the advertised bandwidth is not always reliably available to the customer; ISPs often allow a greater number of subscribers than their backbone connection can handle, under the assumption that most users will not be using their full connection capacity very frequently. This aggregation strategy works more often than not, so users can typically burst to their full bandwidth most of the time; however, peer-to-peer (P2P) file sharing systems, often requiring extended durations of high bandwidth, stress these assumptions, and can cause major problems for ISPs who have excessively overbooked their capacity. For more on this topic, see traffic shaping. As take up for these introductory products increases, Telcos are starting to offer higher bit rate services. For existing connections, this most of the time simply involves reconfiguring the existing equipment at each end of the connection.

As the bandwidth delivered to end users increases, the market expects that video on demand services streamed over the Internet will become more popular, though at the present time such services generally require specialized networks. The data rates on most broadband services still do not suffice to provide good quality video, as MPEG-2 video requires about 6 Mbit/s for good results. Adequate video for some purposes becomes possible at lower data rates, with rates of 768 kbit/s and 384 kbit/s used for some video conferencing applications, and rates as low as 100 kbit/s used for videophones using H.264/MPEG-4 AVC. The MPEG-4 format delivers high-quality video at 2 Mbit/s, at the high end of cable modem and ADSL performance.

Increased bandwidth has already made an impact on newsgroups: postings to groups such as alt.binaries have grown from JPEG files to entire CD and DVD images. According to NTL, the level of traffic on their network increased from a daily inbound news feed of 150 gigabytes of data per day and 1 terabyte of data out each day in 2001 to 500 gigabytes of data inbound and over 4 terabytes out each day in 2002.

Technologies

- Multilinking Modems
- Load Balancing
- ISDN
- T-1/DS-1
- Wired Ethernet

- Rural broadband
- Satellite Internet
- Cellular Broadband
- Remote DSL
- DSL repeater
- Power-line Internet
- Wireless ISP
- iBlast
- WorldSpace

2. Internet Hosting Service

An **Internet hosting service** is a service that runs Internet servers, allowing organizations and individuals to serve content to the Internet. There are various levels of service and various kinds of services offered.

A common kind of hosting is web hosting. Most hosting providers offer a combined variety of services. Web hosting services also offer e-mail hosting service, for example, DNS hosting service is usually bundled with domain name registration.

Web hosting technology has been causing some controversy lately as Web.com claims that it holds patent rights to the hosting technology with its 19 patents. Hostopia, a large wholesale host, recently purchased a license to use that technology from web.com for 10% of retail revenues. Web.com recently sued Go Daddy as well for similar patent infringement.

Generic, yet rather powerful, kinds of Internet hosting provide a server where the clients can run anything they want (including web servers and other servers) and have Internet connections with good upstream bandwidth.

Type of Hosting

1. Full featured hosting
2. Web hosting Service
3. E-mail hosting service
4. DNS hosting service

1. Dynamic DNS

Dynamic DNS is a system which allows the domain name data held in a name server to be updated in real time. The most common use for this is in allowing an Internet domain name to be assigned to a computer with a varying (dynamic) IP address. This makes it possible for other sites on the Internet to establish connections to the machine without needing to track the IP address themselves. A common use is for running server software on a computer that has a dynamic IP address, as is the case with many consumer Internet service providers.

To implement dynamic DNS it is necessary to set the maximum caching time of the domain to an unusually short period (typically a few minutes). This prevents other nodes on the Internet from retaining the old address in their DNS cache, so that they will typically contact the name server of the domain for each new connection.

Dynamic DNS is an integral part of Active Directory, due in part to the fact that domain controllers register their SRV resource records in DNS so that other computers in the Domain (or Forest) can find them.

Types of Dynamic DNS

The term 'dynamic DNS' can be applied to any mechanism for changing a DNS entry from a DNS client. For example, there are many commercial and noncommercial Dynamic DNS providers which provide Dynamic DNS service. Updates by client are done by the **ddclient** program. It sends HTTP GET requests to the Dynamic DNS provider's server which in turn updates the DNS entry.

- Dynamic DNS service is provided on a large scale by various DNS hosting services, which retain the current addresses in a database and provide a "client" program to the user which will send an update to the service whenever the server's IP address has changed. Many routers and other networking components contain a feature such as this in their firmware. The first router to support Dynamic DNS was the UMAX UGate-3000 in 1999, which supported the TZO.COM dynamic DNS service.
- 'Dynamic DNS' can also refer to the protocol documented by RFC 2136 implemented by the **nsupdate** utility. As updating DNS can be dangerous, TSIG can be used to authenticate dynamic DNS updates to a DNS server using HMAC-MD5 hash key. A drawback is that the key needs to be installed on every client in order to use DDNS securely. Microsoft elected to develop an alternative GSS-TSIG, which uses Kerberos for authentication and thus avoids the need for manual installation of hash keys. GSS-TSIG is a proposed standard and is the only authentication supported by Microsoft Windows 2000, Windows XP, and Windows 2003. GSS-TSIG is supported in BIND version 9.5, which previously supported only the TSIG authentication scheme set out in RFC 2845. Microsoft DNS is otherwise fully interoperable with BIND in versions of BIND subsequent to 8.2.2 (when SRV Record support was introduced), and Microsoft has for some time provided information on BIND interoperability with Windows DNS name spaces.

4.6 Leased Lines

A **leased line** is a symmetric telecommunications line connecting two locations. It is sometimes known as a 'Private Circuit' or 'Data Line' in the UK. Unlike traditional PSTN lines it does not have a telephone number, each side of the line being permanently connected to the other. Leased lines can be used for telephone, data or Internet services. Some are ring down services, and some connect two PBXes.

In the U.K., leased lines are usually available at speeds of 64k, 128k, 256k, 512k, 2M and provided to the customer on X.21 presentation. Higher speeds are available on alternative interfaces.

In the U.S., low-speed leased lines (56 kbit/s and below) are usually provided using analog modems. Higher-speed leased lines are usually presented using **FT1 (Fractional T1)**: a **T1** bearer circuit with 1 to 24k 56k or 64k timeslots. Customers must manage their own network termination equipment—Channel Service Unit or Data Service Unit (CSU/DSU).

In Hong Kong, leased lines are usually available at speeds of 64k, 128k, 256k, 512k, T1 (channelized or not) or E1 (less common). Whatever the speed, Telcos usually provide the CSU/DSU and present to the customer on V.35 interface.

For many purposes, leased lines are gradually being replaced by DSL and metro Ethernet.

4.7 Summary

In this unit we learn about Internet Service Providers (ISPs) and their different connection options. It explains Cable Modem, Digital Subscriber Line (DSL), and SHDSL. We know the procedure that how ISP connect to the Internet, and Virtual ISP. Next, Internet hosting services are described with different types of hosting. Finally, broadband access and leased lines are explained briefly.

4.8 Review Questions

1. What do you understand by an Internet service provider?
2. What are different Internet connection options? Name them.
3. Explain typical home-user connection. Write different types of home user connection.
4. Describe DSL and Broadband wireless access. What are different broadband access technologies?
5. Explain cable MODEM.
6. What is a typical business connection? Write different types.
7. Describe SHDSL.
8. Explain how ISP connects to the Internet.
9. What is virtual ISP? Explain.
10. What are different Internet hosting services? Write their types.
11. Explain dynamic DNS. What are different types of dynamic DNS?
12. What is leased line? Explain.



Uttar Pradesh
Rajarshi Tandon Open University

BLISS-03

**Computer Network and
Security Maintenance**

BLOCK

4

Internet Connectivity

Unit 1

Network Operating Systems

5

Unit 2

World Wide Web (WWW) and Client-Server Model

15

Unit 3

Network Planning and Management

27

Unit 4

Network Security

38

Block Introduction:

This is the fourth block of the Data Communication and Networking. In this block different Network Operating Systems are introduced. Next, World Wide Web (WWW) and Client-Server model is described. Aspects of Network planning and management are discussed and lastly concepts of network security are introduced.

In unit 1, there is introduction to different Network Operating Systems. It identifies major network services and describes network protocols associated with major network operating systems. The unit concludes by giving differences among different operating systems.

Unit 2 starts by introducing Client-Server model and defining World Wide Web (WWW). It also describes different client and server configurations.

In unit 3 aspects of network planning and management are introduced. It explains the benefits of network planning and identifies the need for network management. The unit, lastly, lists different hardware devices and software tools used to gather information for network management and planning.

Unit 4 is mainly focused about Network Security. It describes cryptography, encryption, and authentication techniques. Next, it introduces firewalls and proxy servers. Lastly, virtual private networks (VPN) are described briefly.

Unit I – Network Operating Systems

What You Will Learn (Objectives)

After completing this unit you will be able to:

- Introduce network operating systems.
- Identify and describe four major network services.
- Identify network protocols associated with major network operating systems.
- List key differences among network operating systems.

1.1 Introduction

Computers have operating system software that allows them to function. Without operating system instructions, a computer is nothing more than a box with circuits. This is the same with networks. Without a Network Operating System, a network is nothing more than a number of computer devices connected together. In order to transmit information and communicate across a network, it is necessary to have a Network Operating System. There are several different Network Operating Systems, each with its own set of features and protocols. This lesson introduces you to several different Network Operating Systems.

1.2 Terminology

Driver— Also referred to as a device driver. Software that allows communication between the computer and an input/output port or external device. For example, a driver allows a network operating system to communicate with the Network Interface Card.

File Servers— Store files created by application programs.

Network Operating System— A combination of software programs that instruct computers and peripherals to accept requests for services across the network and then provide those services.

Print Server— Accept print jobs sent by anyone across the network.

User Account—An account used by Windows NT Server Operating Systems and other NOS's that provides access to the network. Each user on the network has his/her own unique user name.

User Manager for Domains— A Windows NT Server application program that is used to maintain individual and group user accounts.

Workgroup— Group of devices logically networked together as a single unit. This simplifies network management by segmenting and organizing users into defined groups that can be managed as one. For example, everyone in the accounting department may be considered a workgroup. One change in security for the group changes security for all users.

1.3 Network Operating Systems

In order to transmit signals across a network, it is necessary for the computer to communicate with its modem or Network Interface Card. Network Operating Systems (NOS) provide the protocols necessary to achieve this goal, but each different type of modem or NIC needs to be able to communicate with the particular NOS. It is therefore necessary to install the special software that comes with the interface device. This software is often referred to as a driver. Computers made today usually come with both the interface and necessary drivers installed. Occasionally, you must install the modem or NIC yourself. It is necessary to install the correct driver for that interface device. Failure to so install the driver means that the device will be unable to communicate over the network or with the computer it is installed in.

Network Operating Systems not only allow communication across a network, they also allow a network administrator to organize resources, control access and ensure that the network is operating efficiently. Sharing of network resources can be peer to peer or client server. Which one is the best is dependent on the end goal of the network.

In peer-to-peer networking there is a complete sharing of resources, both hardware and software. All systems act as both users of resources and providers of resources, but no one system is dedicated to a single function. Peer-to-peer networks are generally best suited to small networks and usually are less expensive than client/server networks.

Client/server networks dictate that systems are most often dedicated to a single function. They are either users of network resources or providers of resources. Client/server networks are typically more expensive and robust than peer-to-peer networks and generally support the building of larger networks.

Discussion of Internetworking Operating Systems will cover the four major systems currently in use: Windows, Novell, UNIX/LINUX, and Mac.

1.4 Windows for Workgroups/Windows 95/Windows NT Server

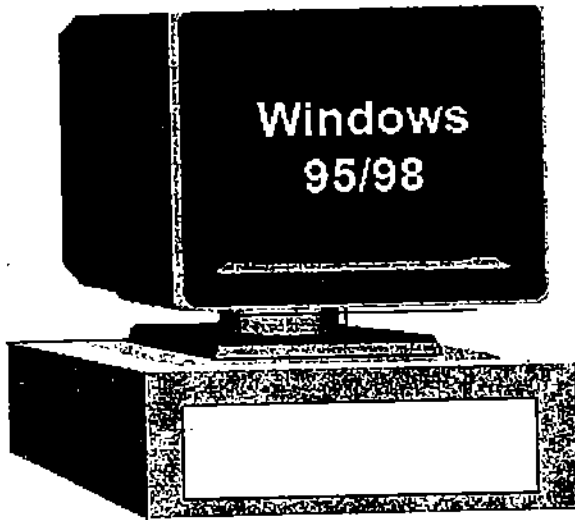
Windows for Workgroups/Windows 95

Windows for Workgroups, introduced in the early 90s and Windows 95, introduced in 1995 are considered peer-to-peer networking systems and do not have the capabilities of true internetworking operating systems. They are, however, inexpensive and more than adequate for small workgroups wanting to share resources, use email, and connect to the Internet.

Windows for Workgroups and Windows 95 both offer peer-to-peer network protocols (Windows NT will be discussed later). The protocols used by these operating systems allow users to share files and devices over LANs. Both offer NetBEUI, Microsoft's small network protocol. They also offer TCP/IP, and IPX/SPX protocols to access the network through either a dial-up connection/modem, or directly through a NIC.

NetBEUI protocols, while not routable, are more than adequate to meet small LAN needs. NetBEUI protocols are easy to use and do not require in-depth networking knowledge. NetBEUI software identifies computer devices by name and it is certainly easier to remember that a computer's name is Juanita or Justin than 141.252.20.2 or 141.252.10.1. Each device name in a network must be unique. NetBEUI software installed on each of the networked computers is all that is necessary to configure devices in order to share resources and create a network. If a small company does want to connect to the Internet, the necessary software and protocols are available with these operating systems.

Windows 95



Shared resources on Windows for Workgroups/95 networks are accessed by a password that protects the resource and there is only one level of access; either you have access or you don't have access. Also user-by-user passwords are not part of the protocols unless Windows NT is present. What this means is that anyone connected to the network who knows the password of the resource has access to that resource. This can create security issues since there is no way to prevent a user from access once s/he knows the password. As the network grows, it is usually more difficult to keep resource passwords secure. Since there is no central control, managing these peer-to-peer networks becomes an issue when the network becomes too large.

1.4.1 Windows NT Server

Windows NT Server was first introduced in the mid-90s and is capable of managing workgroups just like Windows for Workgroups/95. It is a client/server networking operating system that uses routable protocols, making it a true Internetworking Operating System enabling network administrators to connect LANs to WANs. Windows NT Server also provides services for OS/2 and Novell NetWare clients and is able to run on various microprocessors such as Intel Pentium, DEC Alpha, RISC, MIPS, and PowerPC based systems.

Windows NT Server has all of the advantages mentioned for the other Windows operating systems, plus, it contains several other features making it more robust. The security on Windows NT allows a network administrator to not only provide passwords for resources but also to individuals or groups. This operating system does require the use of a more powerful server computer whose sole function is to act as administrator of the NOS program. Having a server where all access data is stored makes managing Windows NT Server Operating Systems efficient.

Windows NT Server has more than one level of security beyond access/no access. This NOS offers:

- no access.
- access that restricts the user to read only capabilities.
- access that allows read and write usage.
- access that allows you to change access permissions for network users.

Each user who wishes to access services on the network must have a password and a user account set up within the domain. A domain is a security model where the database of user accounts is stored on one or more computers known as domain controllers. What this does is centralize control of the network. The network administrator creates, deletes, and manages these accounts and passwords using the User Manager for Domains program that comes with Windows NT Server.

For security reasons, companies often have two servers capable of authenticating passwords, one that acts as the Primary Domain Controller (PDC) and the other as the Backup Domain Controller (BDC). The PDC is a computer on the network that maintains a database of users and security policies for the domain. This database contains the names of users and the resources they may access. User access may be defined for individual files or for entire directories. Note that although the PDC is often a server computer, this is not necessary. Network services can be implemented on one computer and the PDC functions on another.

A BDC maintains a copy of the PDC database. This backup database prevents network disruption in the event of failure of the PDC. If a failure occurs, the backup domain controller can be easily switched to the PDC.

With Windows NT Server networks can connect to other domains. There are four basic domain models:

Single Domain Model—In the single domain model, there is only one defined set of security and user accounts. This means that all management functions are centralized.

Master Domain Model—In this model, there is one master domain server that has the defined set of security and user account data of all other domain servers. However, each of these has only the specific security data for one domain. A user on one domain cannot get permissions not found on her/his domain server. This improves on the single domain model since it allows segmentation of networks, which increases security measures, yet still allows centralized management of network user accounts and accessibility.

Multiple Master Domain Model—The Multiple Master Domain Model has several master domain servers, each with their own specific domains. In this model, network management becomes somewhat decentralized and network administrators cannot maintain the entire network from one central domain server.

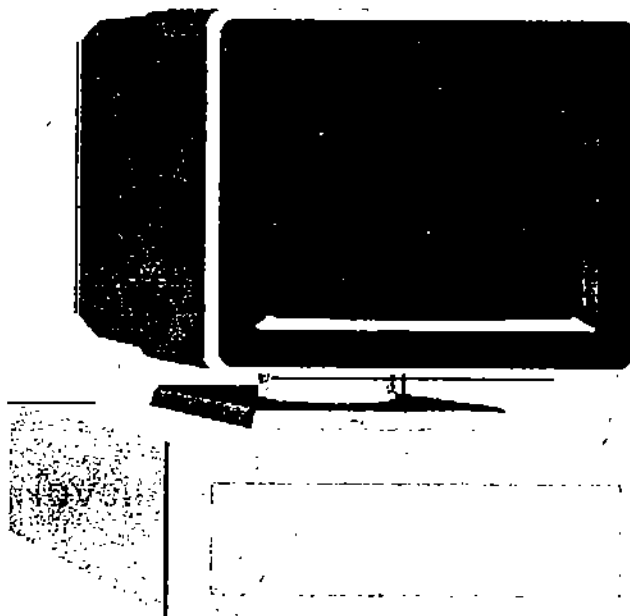
Multiple Trust Domain Model—This model is really a peer-to-peer relationship among domain servers, therefore it becomes decentralized and security is

essentially the same as with Windows for Workgroups and Windows 95.

1.4.2 Novell NetWare

Novell NetWare Operating Systems, also a client/server based NOS, is not domain based but binary based. It has evolved over time from NetWare 2.X, now obsolete, which was designed for small workgroup environments to NetWare 5.X that is aimed at global enterprise network environments. NetWare is optimized for managing, sharing, translating, and synchronizing information throughout the network-computing environment.

Novell NetWare



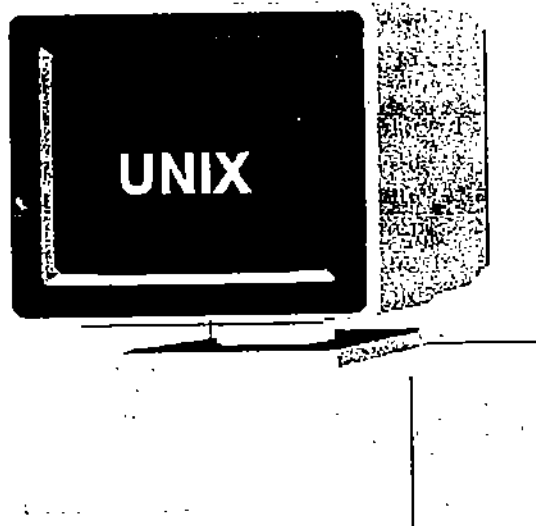
Novell NetWare 4.X features NetWare Directory Services (NDS), which allows a user to logon from anywhere on the network and access the same resources regardless of where the user logs on. This product has become Novell's migration strategy for the future.

Novell NetWare does not actually provide a computer operating system for client workstations; it supplies an operating system for a dedicated server computer, and networking support to existing client workstation operating systems.

7.5 UNIX/LINUX

UNIX is the oldest network operating system still being widely used today. It can be used on either peer-to-peer or client/server networks. LINUX is basically a free version of UNIX, which is developed cooperatively by a community of expert programmers. Various computer manufacturers have released proprietary versions of UNIX which run on their particular hardware platforms. Because of differences in processor hardware and variations from UNIX standards, each company's operating system has its own "flavor". End user applications for UNIX must be individually compiled for each different hardware platform and operating system variant. The networking components in UNIX have been under development for many years, and are based on the collaborative efforts of many developers. Consequently, UNIX networking is extremely reliable. This is why many companies use UNIX to provide their primary networking services and interface to the Internet. In spite of the rise in popularity of Web and Internet services deployed on Windows NT, for security and reliability UNIX is a popular choice in commercial and university environments. Networking under UNIX is based on the TCP/IP protocol, which has only been integrated into the Microsoft Windows network operating systems in recent years. The early work, which established the Internet, was based on UNIX platforms. Most of the TCP/IP services used in UNIX are also commonly used on windows operating system with few exceptions.

UNIX



1. Network File System (NFS)—NFS provides hard disk sharing over TCP/IP networks. It is the primary file and directory sharing protocol used in UNIX. NSF

clients are available for Microsoft Windows operating systems, but are generally used for interoperation with UNIX hosts.

- 2 Remote login services (RLOGIN and TELNET)—UNIX has always been a truly multi-user multi-session operating system. Multitasking of applications and processes is extremely robust; much more so than under Windows NT. In addition many users can run private sessions simultaneously under UNIX, by logging in to a host using a character-based terminal or emulated terminal over a serial or network connection. RLOGIN and TELNET protocols support logging in over network connections.
- 3 Graphical user interface windowing system (X Windows)—X Windows is a completely distributed graphical user interface system. Using X - Windows, a user can execute an application on one computer, and let that application interact with a user on a different computer, using a network connection. X Windows allows computers to share their video displays, keyboards and pointing devices with applications running on other computers. There is no requirement that the application execution and display hosts even have the same type of hardware or run the same operating system, as long as the client application and the display server support the X Windows protocol. Windows NT has recently included support for graphical session-based terminal servers that offer some similar functionality to X Windows, but with much less flexibility. In addition, X Windows server software is available for Microsoft Windows operating systems, allowing applications running on UNIX hosts to be displayed on computers running Microsoft Windows operating systems.

1.6 MAC OS Apple Share

AppleShare provides network services for the Mac OS operating systems. AppleShare supports file and printer sharing over several types of physical networks by using one of the AppleTalk transport protocols: LocalTalk, EtherTalk, TokenTalk or FDDITalk. As the Internet has grown in popularity, TCP/IP software has been developed for the Apple Macintosh computers, along with the standard TCP/IP client applications like Web Browsers and FTP file transfer clients.

AppleShare's peer-to-peer networking is used in small or moderately sized workgroup settings, but has not been adopted widely in large-scale corporate LAN environments.

1.7 Summary

In this unit, we learned the following:

- Identify and describe four major network services.
- Identify network protocols associated with these network operating systems.
- List major differences among local operating systems.
- Compare and contrast peer-to-peer to client/server network operating systems.

- Add a new Operating System to the networked environment.

1.8 Review Questions

- ♦ What are the advantages and disadvantages of Windows for Workgroups/95?
- ♦ What are some of the features added to Windows NT Server that are not supplied for Windows for Workgroups/95?
- ♦ What are some differences between NetWare and Windows NT Server?
- ♦ How has the popularity of the Internet affected network support in Mac OS and the Microsoft Windows operating systems?
- ♦ Which family of operating systems has traditionally supported TCP/IP and internetworking?

Networking Operating Systems

Part A

1. The main function of a Windows NT Domain is to
 - a. Provide shared disk space to a network
 - b. Control user credentials and security access for network from a single location
 - c. Provide shared printer resources to a network
 - d. Connect to the Internet
2. List four major network services

Part B

1. Which protocol is not supported by Windows 95?
 - a. NetBEUI

- b. IPX/SPX
- c. TCP/IP
- d. AppleTalk

2. Which protocol supports remote login sessions?

- a. TELNET
- b. X-Windows
- c. RIP
- d. SPX/IPX

3. Which protocol is not involved with file sharing?

- a) NFS
- b) AppleTalk
- c) NetBEUI
- d) SPX/IPX
- e) X Windows

Part C

1. A network operating system

- a. Supports application development
- b. Supports a single network protocol
- c. Supports resource administration and network communication
- d. Only supports servers
- e. Does not require software drivers

2. Which processor does Windows NT not support?

- a. Intel Pentium
- b. DEC Alpha
- c. IBM RISC
- d. Sun SPARC

Windows 95

- a. Can connect WAN s
- b. Provides centralized security
- c. Is a peer-to-peer Network Operating System
- d. Is too expensive for small workgroups
- e. Can not connect users with the Internet

Part D

1. Peer-to-peer networks

- a. Allow workstations to share and access resources
- b. Are optimized for sharing resources from a single computer with many users
- c. Do not share their own resources
- d. Allow workstations to share their resources with others, but not access others' resources

2. Client/server networks

- a. Share and access resources equally
- b. Are purely database applications
- c. Do not provide access security
- d. Only support a single networking protocol
- e. Support dedicated server computers which are optimized to share resources with many users

Unit II – World Wide Web & Client-Server Model

What You Will Learn (Objectives)

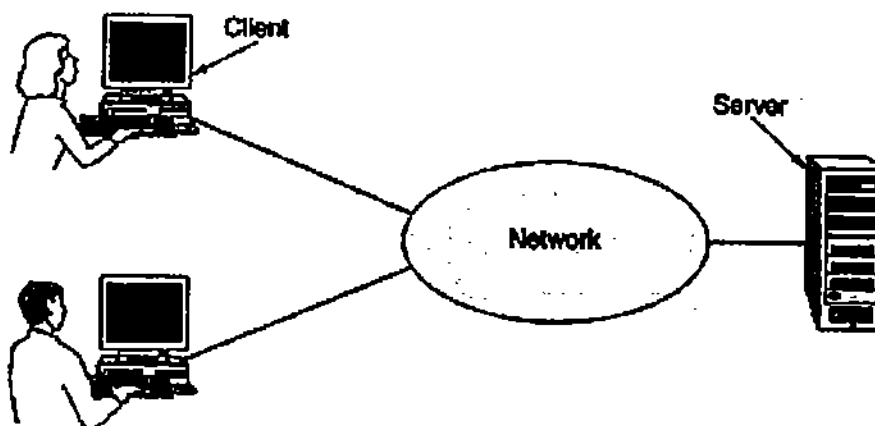
After completing this unit you will be able to:

- Introduce client server model.
- Define World Wide Web (WWW)
- Describe the architecture of WWW
- Understand client configuration
- Understand server configuration

2.1 Introduction

In the simplest of terms, one can imagine a company's information system as consisting of one or more databases and some number of employees who need to access them remotely. In this model, the data are stored on powerful computers called servers. Often these are centrally housed and maintained by a system administrator. In contrast, the employees have simpler machines, called clients, on their desks, with which they access remote data, for example, to include in spreadsheets they are constructing. (Sometimes we will refer to the human user of the clients machine as the "client," but it should be clear from the context whether we mean the computer or its user.) The clients and server machines are connected by a network, as illustrated in figure. Note that we have shown the network as a simple oval, without any detail. We will use this form when we mean a network in the abstract sense. When more detail is required, it will be provided.

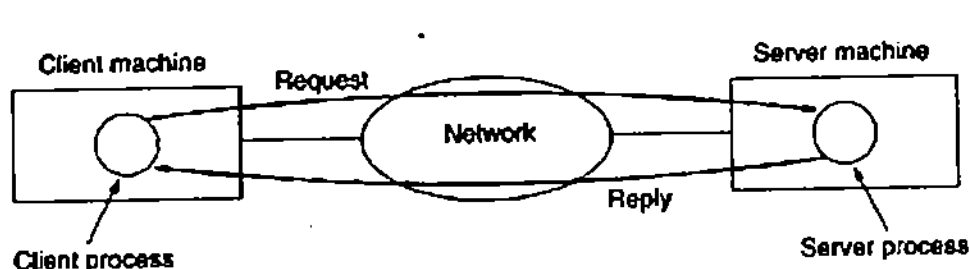
Figure: A network with two clients and one server



This whole arrangement is called the **client-server model**. It is widely used and forms the basis of much network usage. It is applicable when the client and server are both in the same building (e.g., belong to the same company), but also when they are far apart. For example, when a person at home accesses a page on the World Wide Web, the same model is employed, with the remote Web server being the server and the user's personal computer being the client. Under most conditions, one server can handle a large number of clients.

If we look at the client-server model in detail, we see that two processes are involved, one on the client machine and one on the server machine. Communication takes the form of the client process sending a message over the network to the server process. The client process then waits for a reply message. When the server process gets the request, it performs the requested work or looks up the requested data and sends back a reply. These messages are shown in Figure.

Figure: the client-server model involves request and reply



2.2 The World Wide Web (WWW)

The World Wide Web is an architectural framework for accessing linked documents spread out over millions of machines all over the Internet. In 10 years, it went from being a way to distribute high-energy physics data to the application that millions of people think of as being "The Internet." Its enormous popularity stems from the fact that it has a colorful graphical interface that is easy for beginners to use, and it provides an enormous wealth of information on almost every conceivable subject, from aardvarks to Zulus.

The Web (also known as WWW) began in 1989 at CERN, the European center for nuclear research. CERN has several accelerators at which large teams of scientists from the participating European countries carry out research in particle physics. These teams often have members from half a dozen or more countries. Most experiments are highly complex and require years of advance planning and equipment construction. The Web grew out of the need to have these large teams of internationally dispersed researchers collaborate using a constantly changing collection of reports, blueprints, drawings, photos, and other documents.

2.3 Architectural Overview

From the users' point of view, the Web consists of a vast, worldwide collection of documents or Web pages, often just called pages for short. Each page may contain links

to other pages anywhere in the world. Users can follow a link by clicking on it, which then takes them to the page pointed to. This process can be repeated indefinitely. The idea of having one page point to another, now called hypertext, was invented by a visionary M.I.T. professor of electrical engineering, Vannevar Bush, in 1945, long before the Internet was invented.

Pages are viewed with a program called a browser, of which Internet Explorer and Netscape Navigator are two popular ones. The browser fetches the page requested, interprets the text and formatting commands on it, and displays the page, properly formatted, on the screen. An example is given in figure. Like many Web pages, this one starts with a title, contains some information, and ends with the e-mail address of the page's maintainer. Strings of text that are links to other pages, called hyperlinks, are often highlighted, by underlining, displaying them in a special color, or both. To follow a link, the user places the mouse cursor on the highlighted area, which causes the cursor to change, and clicks on it. Although non graphical browsers, such as Lynx, exist, they are not as popular as graphical browsers, so we will concentrate on the latter. Voice-based browsers are also being developed.

Figure: A Web page

WELCOME TO THE UNIVERSITY OF EAST PODUNK'S WWW HOME PAGE

- Campus Information
 - ▣ Admissions Information
 - ▣ Campus map
 - ▣ Directions to campus
 - ▣ The UEP student body
- Academic Departments
 - ▣ Department of Animal Psychology
 - ▣ Department of Alternative Studies
 - ▣ Department of Microbiotic Cooking
 - ▣ Department of Nontraditional Studies
 - ▣ Department of Traditional Studies

Webmaster@eastpodunk.edu

The basic model of how the Web works is shown in Figure. Here the browser is displaying a Web page on the client machine. When the user clicks on a line of text that is linked to a page on the abcd.com server, the browser follows the hyperlink by sending a message to the abcd.com server asking it for the page. When the page arrives, it is displayed. If this page contains a hyperlink to a page on the xyz.com server that is clicked on, the browser then sends a request to that machine for the page, and so on indefinitely.

2.3.1 The Client Side

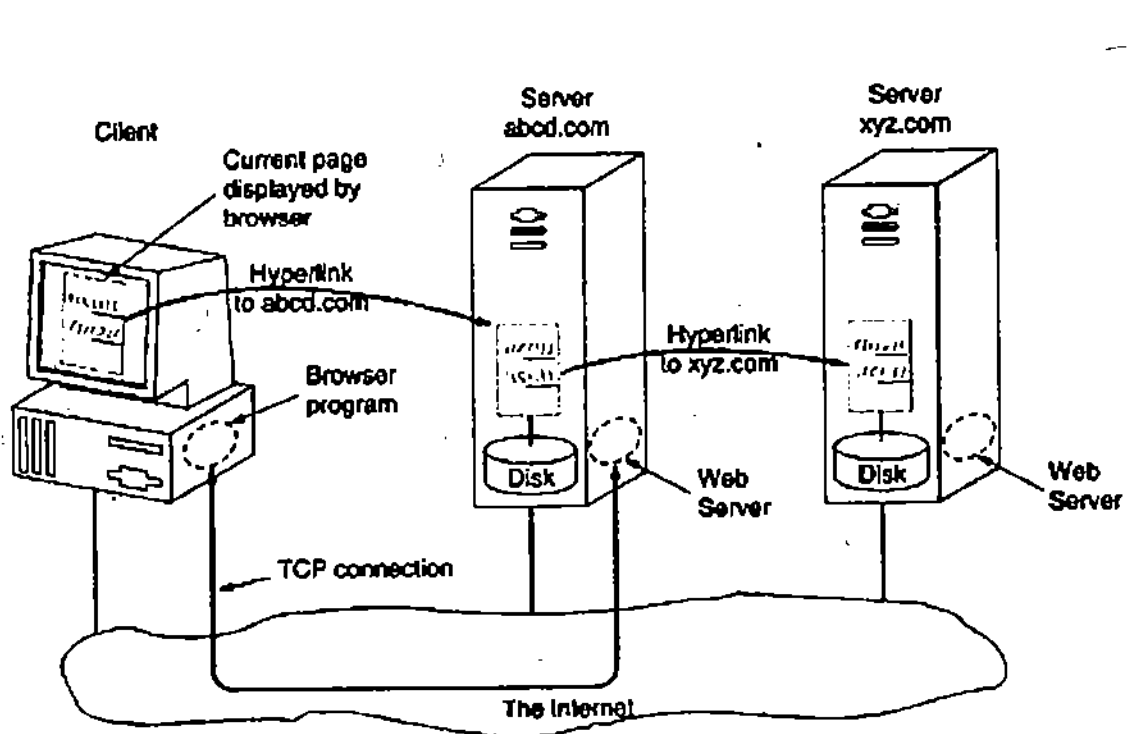
Let us now examine the client side of figure in more detail. In essence, a browser is a program that can display a Web page and catch mouse clicks to items on the displayed page. When an item is selected, the browser follows the hyperlink and fetches the page selected. Therefore, the embedded hyperlink needs a way to name any other page on the Web. Pages are named using URLs (Uniform Resource Locators). A typical URL is

`http://www.abcd.com/products.html`

We will explain URLs later in this chapter. For the moment, it is sufficient to know that a URL has three parts: the name of the protocol (`http`), the DNS name of the machine where the page is located (`www.abcd.com`), and (usually) the name of the file containing the page (`products.html`).

When a user clicks on a hyperlink, the browser carries out a series of steps in order to fetch the page pointed to. Suppose that a user is browsing the Web and finds a link on Internet telephony that points to ITU's home page, which is `http://www.itu.org/home/index.html`. Let us trace the steps that occur when this link is selected.

Figure: The Parts of the Web Model



1. The browser determines the URL (by seeing what was selected).

2. The browser asks DNS for the IP address of `www.itu.org`.
3. DNS replies with `156.106.192.32`.
4. The browser makes a TCP connection to port 80 on `156.106.192.32`.
5. It then sends over a request asking for file `/home/index.html`.
6. The `www.itu.org` server sends the file `/home/index.html`.
7. The TCP connection is released.
8. The browser displays all the text in `/home/index.html`.
9. The browser fetches and displays all images in this file.

Many browsers display which step they are currently executing in a status line at the bottom of the screen. In this way, when the performance is poor, the user can see if it is due to DNS not responding, the server not responding, or simply network congestion during page transmission.

To be able to display the new page (or any page), the browser has to understand its format. To allow all browsers to understand all Web pages, Web pages are written in a standardized language called HTML, which describes Web pages. We will discuss it in detail later in this chapter.

Although a browser is basically an HTML interpreter, most browsers have numerous buttons and features to make it easier to navigate the Web. Most have a button for going back to the previous page, a button for going forward to the next page (only operative after the user has gone back from it), and a button for going straight to the user's own start page. Most browsers have a button or menu item to set a bookmark on a given page and another one to display the list of bookmarks, making it possible to revisit any of them with only a few mouse clicks. Pages can also be saved to disk or printed. Numerous options are generally available for controlling the screen layout and setting various user preferences.

In addition to having ordinary text (not underlined) and hypertext (underlined), Web pages can also contain icons, line drawings, maps, and photographs. Each of these can (optionally) be linked to another page. Clicking on one of these elements causes the browser to fetch the linked page and display it on the screen, the same as clicking on text. With images such as photos and maps, which page is fetched next may depend on what part of the image was clicked on.

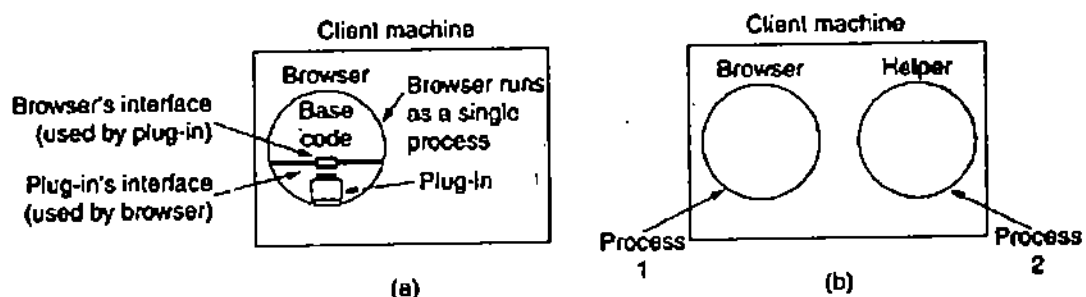
Not all pages contain HTML. A page may consist of a formatted document in PDF format, an icon in GIF format, a photograph in JPEG format, a song in MP3 format, a video in MPEG format, or any one of hundreds of other file types. Since standard HTML pages may link to any of these, the browser has a problem when it encounters a page it cannot interpret.

Rather than making the browsers larger and larger by building in interpreters for a rapidly growing collection of file types, most browsers have chosen a more general solution. When a server returns a page, it also returns some additional information about the page. This information includes the MIME type of the page. Pages of type `text/html` are just displayed directly, as are pages in a few other built-in types. If the MIME type is not one

of the built-in ones, the browser consults its table of MIME types to tell it how to display the page. This table associates a MIME type with a viewer.

There are two possibilities: plug-ins and helper applications. A plug-in is a code module that the browser fetches from a special directory on the disk and installs as an extension to itself, as illustrated in figure. Because plug-ins run inside the browser, they have access to the current page and can modify its appearance. After the plug-in has done its job (usually after the user has moved to a different Web page), the plug-in is removed from the browser's memory.

Figure: (a) A browser plug-in. (b) A helper application



Each browser has a set of procedures that all plug-ins must implement so the browser can call the plug-in. For example, there is typically a procedure the browser's base code calls to supply the plug-in with data to display. This set of procedures is the plug-in's interface and is browser specific.

In addition, the browser makes a set of its own procedures available to the plug-in, to provide services to plug-ins. Typical procedures in the browser interface are for allocating and freeing memory, displaying a message on the browser's status line, and querying the browser about parameters.

Before a plug-in can be used, it must be installed. The usual installation procedure is for the user to go to the plug-in's Web site and download an installation file. On Windows, this is typically a self-extracting zip file with extension .exe. When the file is double clicked, a little program attached to the front of the zip file is executed. This program unzips the plug-in and copies it to the browser's plug-in directory. Then it makes the appropriate calls to register the plug-in's MIME type and to associate the plug-in with it. On UNIX, the installer is often a shell script that handles the copying and registration.

The other way to extend a browser is to use a helper application. This is a complete program, running as a separate process. It is illustrated in figure. Since the helper is a separate program, it offers no interface to the browser and makes no use of browser services. Instead, it usually just accepts the name of a scratch file where the content file has been stored, opens the file, and displays the contents. Typically, helpers are large programs that exist independently of the browser, such as Adobe's Acrobat Reader for

displaying PDF files or Microsoft Word. Some programs (such as Acrobat) have a plug-in that invokes the helper itself.

Many helper applications use the MIME type application. A considerable number of subtypes have been defined, for example, application/pdf for PDF files and application/msword for Word files. In this way, a URL can point directly to a PDF or Word file, and when the user clicks on it, Acrobat or Word is automatically started and handed the name of a scratch file containing the content to be displayed. Consequently, browsers can be configured to handle a virtually unlimited number of document types with no changes to the browser. Modern Web servers are often configured with hundreds of type/subtype combinations and new ones are often added every time a new program is installed.

Helper applications are not restricted to using the application MIME type. Adobe Photoshop uses image/x-photoshop and RealOne Player is capable of handling audio/mp3, for example.

On Windows, when a program is installed on the computer, it registers the MIME types it wants to handle. This mechanism leads to conflict when multiple viewers are available for some subtype, such as video/mpg. What happens is that the last program to register overwrites existing (MIME type, helper application) associations, capturing the type for itself. As a consequence, installing a new program may change the way a browser handles existing types.

On UNIX, this registration process is generally not automatic. The user must manually update certain configuration files. This approach leads to more work but fewer surprises.

Browsers can also open local files, rather than fetching them from remote Web servers. Since local files do not have MIME types, the browser needs some way to determine which plug-in or helper to use for types other than its built-in types such as text/html and image/jpeg. To handle local files, helpers can be associated with a file extension as well as with a MIME type. With the standard configuration, opening foo.pdf will open it in Acrobat and opening bar.doc will open it in Word. Some browsers use the MIME type, the file extension, and even information taken from the file itself to guess the MIME type. In particular, Internet Explorer relies more heavily on the file extension than on the MIME type when it can.

Here, too, conflicts can arise since many programs are willing, in fact, eager, to handle, say, .mpg. During installation, programs intended for professionals often display checkboxes for the MIME types and extensions they are prepared to handle to allow the user to select the appropriate ones and thus not overwrite existing associations by accident. Programs aimed at the consumer market assume that the user does not have a clue what a MIME type is and simply grab everything they can without regard to what previously installed programs have done.

The ability to extend the browser with a large number of new types is convenient but can also lead to trouble. When Internet Explorer fetches a file with extension exe, it realizes that this file is an executable program and therefore has no helper. The obvious action is

to run the program. However, this could be an enormous security hole. All a malicious Web site has to do is to produce a Web page with pictures of, say, movie stars or sports heroes, all of which are linked to a virus. A single click on a picture then causes an unknown and potentially hostile executable program to be fetched and run on the user's machine. To prevent unwanted guests like this, Internet Explorer can be configured to be selective about running unknown programs automatically, but not all users understand how to manage the configuration.

On UNIX an analogous problem can exist with shell scripts, but that requires the user to consciously install the shell as a helper. Fortunately, this installation is sufficiently complicated that nobody could possibly do it by accident (and few people can even do it intentionally).

2.3.2 The Server Side

So much for the client side. Now let us take a look at the server side. As we saw above, when the user types in a URL or clicks on a line of hypertext, the browser parses the URL and interprets the part between `http://` and the next slash as a DNS name to look up. Armed with the IP address of the server, the browser establishes a TCP connection to port 80 on that server. Then it sends over a command containing the rest of the URL, which is the name of a file on that server. The server then returns the file for the browser to display.

The server, like a real Web server, is given the name of a file to look up and return. In both cases, the steps that the server performs in its main loop are:

1. Accept a TCP connection from a client (a browser).
2. Get the name of the file requested.
3. Get the file (from disk).
4. Return the file to the client.
5. Release the TCP connection.

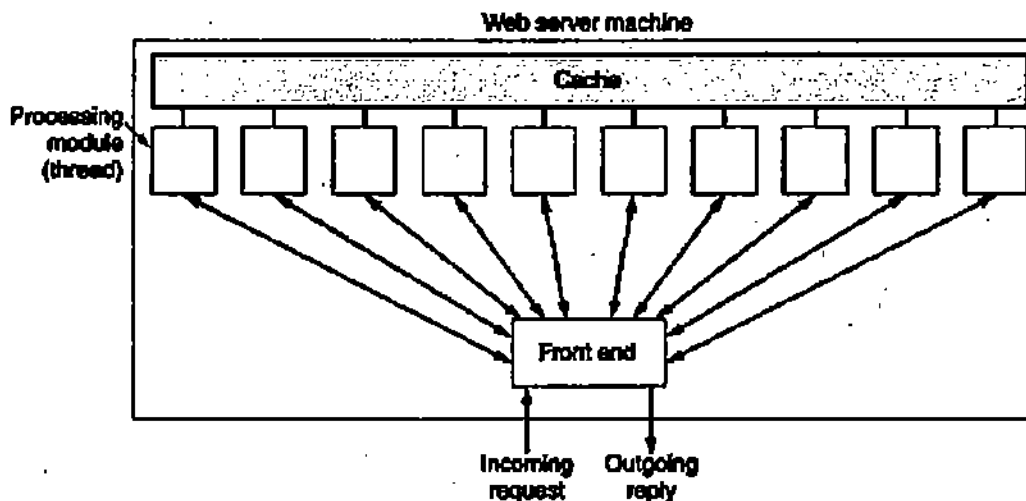
Modern Web servers have more features, but in essence, this is what a Web server does.

A problem with this design is that every request requires making a disk access to get the file. The result is that the Web server cannot serve more requests per second than it can make disk accesses. A high-end SCSI disk has an average access time of around 5 msec, which limits the server to at most 200 requests/sec, less if large files have to be read often. For a major Web site, this figure is too low.

One obvious improvement (used by all Web servers) is to maintain a cache in memory of the n most recently used files. Before going to disk to get a file, the server checks the cache. If the file is there, it can be served directly from memory, thus eliminating the disk access. Although effective caching requires a large amount of main memory and some extra processing time to check the cache and manage its contents, the savings in time are nearly always worth the overhead and expense.

The next step for building a faster server is to make the server multithreaded. In one design, the server consists of a front-end module that accepts all incoming requests and k processing modules, as shown in Figure. The $k + 1$ threads all belong to the same process so the processing modules all have access to the cache within the process' address space. When a request comes in, the front end accepts it and builds a short record describing it. It then hands the record to one of the processing modules. In another possible design, the front end is eliminated and each processing module tries to acquire its own requests, but a locking protocol is then required to prevent conflicts.

Figure: A multithreaded Web server with a front end and processing modules



The processing module first checks the cache to see if the file needed is there. If so, it updates the record to include a pointer to the file in the record. If it is not there, the processing module starts a disk operation to read it into the cache (possibly discarding some other cached files to make room for it). When the file comes in from the disk, it is put in the cache and also sent back to the client.

The advantage of this scheme is that while one or more processing modules are blocked waiting for a disk operation to complete (and thus consuming no CPU time), other modules can be actively working on other requests. Of course, to get any real improvement over the single-threaded model, it is necessary to have multiple disks, so more than one disk can be busy at the same time. With k processing modules and k disks, the throughput can be as much as k times higher than with a single-threaded server and one disk.

In theory, a single-threaded server and k disks could also gain a factor of k , but the code and administration are far more complicated since normal blocking READ system calls cannot be used to access the disk. With a multithreaded server, they can be used since then a READ blocks only the thread that made the call, not the entire process.

Modern Web servers do more than just accept file names and return files. In fact, the actual processing of each request can get quite complicated. For this reason, in many servers each processing module performs a series of steps. The front end passes each incoming request to the first available module, which then carries it out using some subset of the following steps, depending on which ones are needed for that particular request.

1. Resolve the name of the Web page requested.
2. Authenticate the client.
3. Perform access control on the client.
4. Perform access control on the Web page.
5. Check the cache.
6. Fetch the requested page from disk.
7. Determine the MIME type to include in the response.
8. Take care of miscellaneous odds and ends.
9. Return the reply to the client.
10. Make an entry in the server log.

Step 1 is needed because the incoming request may not contain the actual name of the file as a literal string. For example, consider the URL `http://www.cs.vu.nl`, which has an empty file name. It has to be expanded to some default file name. Also, modern browsers can specify the user's default language (e.g., Italian or English), which makes it possible for the server to select a Web page in that language, if available. In general, name expansion is not quite so trivial as it might at first appear, due to a variety of conventions about file naming.

Step 2 consists of verifying the client's identity. This step is needed for pages that are not available to the general public. We will discuss one way of doing this later in this chapter.

Step 3 checks to see if there are restrictions on whether the request may be satisfied given the client's identity and location. Step 4 checks to see if there are any access restrictions associated with the page itself. If a certain file (e.g., `.htaccess`) is present in the directory where the desired page is located, it may restrict access to the file to particular domains, for example, only users from inside the company.

Steps 5 and 6 involve getting the page. Step 6 needs to be able to handle multiple disk reads at the same time.

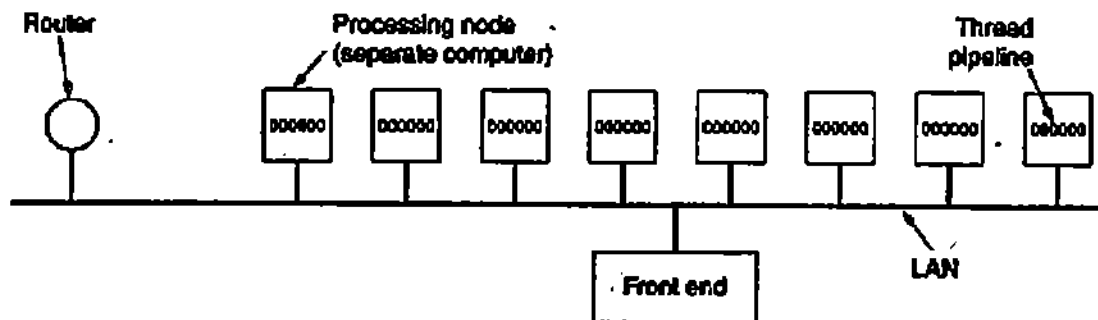
Step 7 is about determining the MIME type from the file extension, first few words of the file, a configuration file, and possibly other sources. Step 8 is for a variety of miscellaneous tasks, such as building a user profile or gathering certain statistics.

Step 9 is where the result is sent back and step 10 makes an entry in the system log for administrative purposes. Such logs can later be mined for valuable information about user behavior, for example, the order in which people access the pages.

If too many requests come in each second, the CPU will not be able to handle the processing load, no matter how many disks are used in parallel. The solution is to add

more nodes (computers), possibly with replicated disks to avoid having the disks become the next bottleneck. This leads to the server farm model of Figure. A front end still accepts incoming requests but sprays them over multiple CPUs rather than multiple threads to reduce the load on each computer. The individual machines may themselves be multithreaded and pipelined as before.

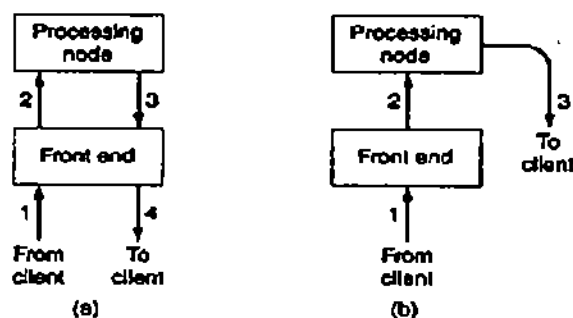
Figure: A Server Farm



One problem with server farms is that there is no longer a shared cache because each processing node has its own memory—unless an expensive shared-memory multiprocessor is used. One way to counter this performance loss is to have a front end keep track of where it sends each request and send subsequent requests for the same page to the same node. Doing this makes each node a specialist in certain pages so that cache space is not wasted by having every file in every cache.

Another problem with server farms is that the client's TCP connection terminates at the front end; so the reply must go through the front end. This situation is depicted in Figure, where the incoming request (1) and outgoing reply (4) both pass through the front end. Sometimes a trick, called TCP handoff, is used to get around this problem. With this trick, the TCP end point is passed to the processing node so it can reply directly to the client, shown as (3) in figure. This handoff is done in a way that is transparent to the client.

Figure: (a) Normal request-reply message sequence. (b) Sequence when TCP handoff used



2.9 Summary

In this unit we learn about:

- Client-Server Model
- World Wide Web (WWW)
- Architecture of World Wide Web
- The parts of WWW
- Client configuration
- Server configuration

2.10 Review Questions

1. Explain client-server model using a diagram.
2. What is World Wide Web (WWW)?
3. Describe the architecture of WWW.
4. Using neat diagram explain client server configuration.
5. Using diagram explain the parts of Web model.

Unit III – Network Planning & Management

What You Will Learn (Objectives)

After completing this unit you will be able to:

- List and define criteria involved with network planning and management
- Explain the benefits of network planning
- Identify the need for network management
- Identify hardware devices and software tools used to gather information for network management and planning

3.1 INTRODUCTION

This unit presents an overview of network planning and management tasks. It also discusses hardware and software tools, and defines terms associated with network planning and management.

Terminology

- **Quality of Service Analysis (QoS)**—Quality of service analysis, which uses management logs, and a model of how objects should perform are important components of this function.
- **Propagation Delay**—Propagation delay measures the time between events: the transmission of a data unit and the arrival of a collision notification from the remotest device.
- **Resource Waiting Time**—Resource waiting time measures the performance of queue lengths, time waiting for a service, time it takes to complete the service.
- **Response Time**—Response time is the time between the sending of a request and the acknowledgment or confirmation of receipt.
- **Throughput**—Throughput measures the network data that is successfully transmitted without errors. The throughput is measured in both directions, to and from a device - traffic that enters the network through the physical medium as well as traffic that leaves the network through the medium.
- **Workload**—Workload specifies requirements for measuring overload conditions and notification of such conditions. Warnings are set to indicate that additional workload will be rejected. Conditions for clearing the rejection warning must also be defined.

3.2 Network Maintenance and Management

There are five major elements to the maintenance and management of a network. They include:

- Fault tolerance management
- Configuration management
- Performance management
- Security management
- Accounting management

The table below summarizes the tasks and managing objects used for each of these five areas.

3.2.1 Fault Management

When networks run mission-critical applications, quick resolution of problems is imperative. When a factory that has an automated manufacturing process "goes down" it can impact other areas of the company. Orders might not be filled, payroll might be late, new products may not get to the market ahead of the competition, and so on. If a hospital network goes down, the downtime can mean failure of a life-saving operation.

Resolving network problems, or faults, is the job of the network manager. Fault management involves several steps. The network administrator must:

- Identify the problem
- Find ways to isolate the cause
- Correct the problem
- Prevent it from occurring again

The focus of fault management is on the network devices and the physical and virtual connections that operate at the three lower layers of the OSI reference model.

Identifying Network Problems

Network problems range from the network being down and completely unavailable to the slowing down of one device, such as a printer. Problems can include the following:

- Hardware errors:

Stopped working

Will not start

Has slowed down

Produces intermittent errors

- Software errors:

Quits unexpectedly

Does not load

Does not work as described in the documentation

Does not work with other applications

Produces unexpected errors

- Data transmission errors:

Connection failure

Collisions

Congestion

Transmission errors

The distinction between hardware, software, and transmission problems is often indistinguishable and requires that system tools be constantly updated to keep up with networking improvements.

Device Isolation

If a user's workstation fails, troubleshooting is a relatively simple procedure. Among the procedures/tools that aid in isolating a fault when an operational device suddenly fails include the following:

- Checking for loose connections
- Using diagnostic procedures that have been built into the device
- Use loopback tests to determine whether the correct logical connections exist
- On a TCP/IP network, use the ping command and netstat utility for basic troubleshooting assistance:

The ping command determines whether a device is connected and operating over the network.

- The netstat utility shows the active device drivers on the interface cards, the number of datagrams sent and received, the routers that a given workstation has "learned" about, and the connections that are established.

Software Fault Isolation

If the hardware works, then software can be checked for:

- Configuration errors

- Conflicts with other applications in use

General Failure

In the case of a general failure, tools that check or monitor network activity and provide reports on network errors help determine and isolate the problem. If a device has not failed completely, it may have sent a message reporting that it was shutting down, or that a number of transmission errors had occurred.

If a device has failed without sending out a message, each network device may need to be checked in turn. These devices can be any of the following:

- A hub through which user workstations connect to the network
- A server where user applications and/or files are stored
- An internetworking device that connects one network to another

Data Transmission

Data transmission errors can be reported by monitoring devices. Protocol analyzers capture and examine actual data, showing which devices have been communicating, what protocols are actively in use and which transmissions have been experiencing errors.

Fault Management Tools

Fault management tools provide the network manager with information that helps to isolate problems and can be used for troubleshooting purposes, such as:

- Event reporting—Systems can be configured to report events such as the shutting down of a device, activation, or deactivation of a connection.
- Diagnostic testing—Diagnostic tools can determine the integrity of a cable, locate overloaded network devices, or diagnose protocol problems.
- Alarm reporting—devices can send an "alarm alert" to a monitoring device when given events occur or when thresholds are reached (the network is operating at 60% capacity, the storage device is 80% full).

Configuration Management

When adding a device to a small network, the installer can often use the default values configured by the vendor. Generally, default settings can also be used for software installed in small networks. As the users become more sophisticated and networks more complex, configurations must be customized in order to accommodate connectivity and interoperability and issues.

When devices and software work together well, configurations of network resources require no modification. However, the network manager needs to continuously monitor the state of the network to ascertain that the network is

operating as desired. When needed, the network manager can make changes in logical and physical configurations for more effective network operation.

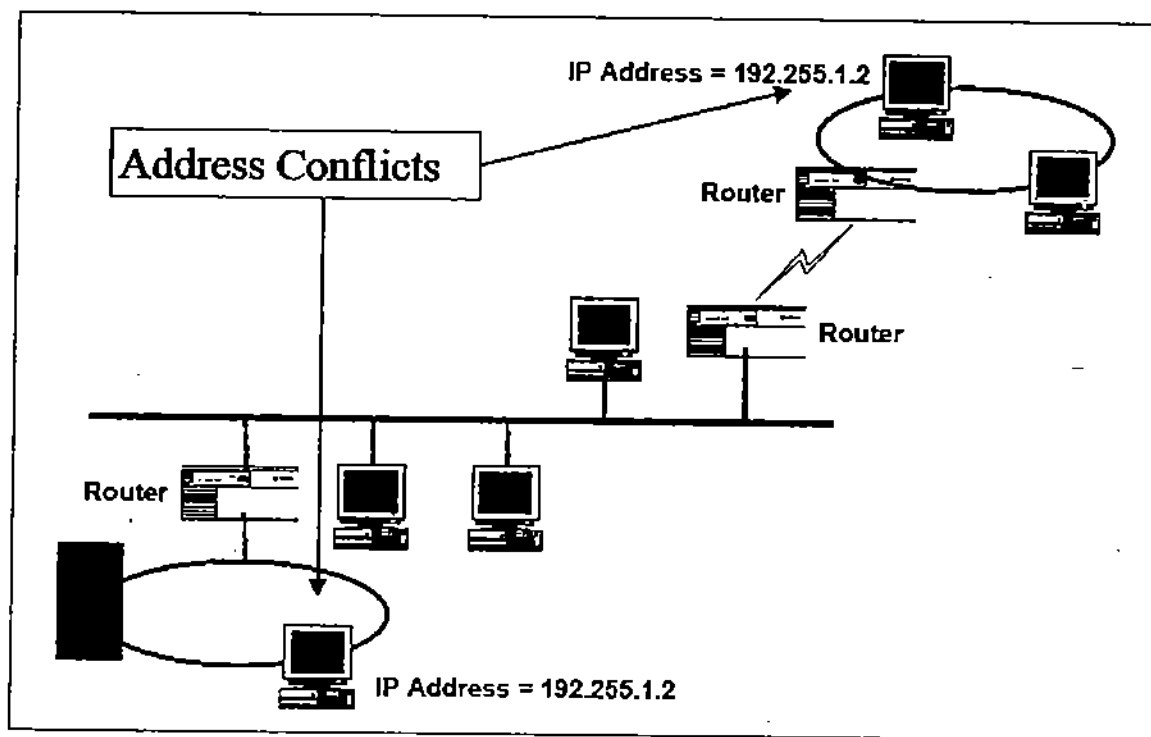
Modifying Configurations

The configuration of devices may need modification for the following reasons:

- Network changes
- Availability

Changes in the network may suddenly cause problems, making some vital aspect of the network malfunction. For example, a network with an unregistered IP address may now want to connect to the internetwork. Unless already configured, all IP addresses for devices on a network must be unique. The following graphic illustrates an address conflict that has occurred after a network has connected to an internetwork.

IP address conflicts



When a second file server is installed on a small network, the configuration of the first server may need to be changed so both servers can be used efficiently.

A device initially available to one group only may now need to be more generally accessible on the network. Access to files in a sales office database or on a mainframe computer in the main office may be limited to one site or even one group in the organization. As the organization grows, people in the new site might request access to these resources.

Lack of Use

Often the addition of a faster connection between two sites, or more efficient devices might cause a decrease in the use of the slower connection. It may be more cost-effective to eliminate the connection or move it to another workgroup. It may need to be discarded completely.

Objectives/Polices

Organizational objectives, such as the adoption of devices using the latest technology, can result in frequent configuration changes as these devices are introduced to the network.

Configuration Management

Responsibilities associated with managing configurations include the following:

- Developing a configuration map of the devices on the network. Having this map will help the network manager anticipate modifications needed when a new connection is added. If devices have been reconfigured from their default settings, that should be noted.
- Developing a configuration map of the software on the network. The configuration information should include memory requirements, memory buffers, driver settings and interoperability issues. A map will also help track and isolate problems.
- Adding devices to the configuration map so that it is current. Note which devices it is connected to and indicate any issues or concerns that should be monitored.
- Ensure that only one version of software is operating on the network and that all users are notified of version changes.
- Change the operational characteristics of managed objects and record any changes in the state of the objects.
 - a. Enabled: Not in use, but operational and available
 - b. Disabled: Not available, or dependent on another source that is not available
 - c. Active: Available and can accept services
 - d. Busy: Available, but cannot accept additional services

3.2.2 Performance Management

Performance management is a continuous monitoring and tuning process designed to:

- Measure system performance
- Determine if performance is satisfactory
- Adjust measurement criteria
- Plan for future needs

Establishing a Benchmark

Performance management starts with establishing a benchmark, or baseline, to determine the current network performance. For example, to collect baseline data, and keep track of data traffic and overall network use:

- Record any physical change made to the network on the configuration map previously prepared (discussed under configuration management)
 - Determine the baseline (normal, operational) performance of the network: track daily use, variation over time, and occurrence of different types of low-level errors
 - Keep a record of actual network use, which is the number of bits transmitted over the network
 - Examine transmission errors, such as frames that are too long, collisions, and runt frames (frames shorter than the minimum allowable)
 - Record the types of protocols used on the network
 - Eliminate bottlenecks to improve resource availability
- Once a baseline is established, monitor the network for deviations.

Performance Measurements

The OSI management standard defines the following types of performance measurements:

- Workload
- Throughput
- Resource waiting time
- Response time
- Propagation delay

• Any quality of service (QoS) change

Workload specifies requirements for measuring overload conditions and notification of such conditions. Warnings are set to indicate that additional

workload will be rejected. Conditions for clearing the rejection warning must also be defined.

Throughput measures the network data that is successfully transmitted without errors. The throughput is measured in both directions, to and from a device - traffic that enters the network through the physical medium as well as traffic that leaves the network through the medium.

Resource waiting time measures the performance of queue lengths, time waiting for a service, time it takes to complete the service.

Response time is the time between the sending of a request and the acknowledgment or confirmation of receipt.

Propagation delay measures the time between events: the transmission of a data unit and the arrival of a collision notification from the remotest device.

Quality of service analysis, which uses management logs, and a model of how objects should perform are important components of this function.

3.2.3 Security Management

The security measures required by a network depend upon:

- Sensitivity of the information
- Levels of access that are needed

A basic security measure is password assignment. Passwords limit access to network resources as follows:

- Users have passwords so that only authorized users can log on to the network.
- Passwords can also be assigned to software, hardware devices, and to directories on the network.

Passwords can be encrypted so that they cannot be easily decoded as they pass through the network system.

Audit trails and alarm systems (notification of unauthorized attempt to access a resource) add to these password security measures.

In addition to password security, rights to network resources can be assigned so that some users can have selective levels of access. For example, the network administrator and the person who maintains a database might have read-write rights, so that they can add, modify, or copy data. Other users can be granted rights to copy information from the database, but not to modify the data. Still others might be able to add to the database, but not copy or modify the data. Finally, other users might be completely excluded from the use of the database.

Additional services identified by OSI documents on security management, but not yet defined, include those related to:

- Data integrity
- Confirm that the source of the data is as stated.

- Confirm that the data has not been changed in transit.
- Protect data from modification, deletion, or analysis. Services can ensure that data recovery procedures are executed or allow the network manager to specify that no recovery procedure can be attempted.
- Delivery of information
- Guarantee accuracy, origin, timing, and delivery of information.
- Select data paths that bypass certain networks for purposes of security through routing controls.
- Control who can block the sending or receipt of information.

3.2.4 Accounting Management

In small systems, the network manager may want to monitor the use of network disk storage space or network activity. Statistics about network use provides valuable data for performance, configuration, and fault management. In addition, the data can be used for tracking access time and billing for access time.

Accounting specifications have not been completed by OSI. Rules under consideration include the use of accounting meters, which are triggers for updating data and for reporting usage.

3.3 Network Management Tools

A variety of hardware and software tools are available to support network management. SNMP compliant devices like routers and switches can transmit information to be collected by SNMP software applications. These applications can report configuration information, status or faults for a wide array of measurements. SNMP applications can also be used to control network devices and to change general network configuration. Many network devices include simple diagnostic displays on their front panels. Even network interface cards usually have some sort of diagnostic readout that can help diagnose status and fault information. Packet sniffing is useful for detecting problems in network transmission, which may be either hardware or software related faults. Packet sniffer software applications like Sniffer Basic must be executed on a computer that is physically connected to the network. Some hardware sniffers are basically hand-held computers with network interface cards that must be connected to the network to capture packets. Others use inductive interfaces to capture packets without being directly connected to the network cable.

3.4 Summary

In this unit, we learned the following:

- Criteria involved with network planning and management
- The benefits of network planning
- How to identify the need for network management
- Identify hardware devices and software tools used to gather information for network management and planning

3.5 Review Questions

- ♦ What are the five major areas of network management?
- ♦ Name three types of network problems:
- ♦ List areas for which the OSI has defined standards.
- ♦ What might a network administrator do to increase security management?
- ♦ Why is it important for a network administrator to establish benchmarks?

Network Maintenance and Management

Part A

1. What are the four steps in the fault management process?
2. Name several hardware, software, and data transmission errors.
3. What are the five major elements to the maintenance and management of a network?

4. Network planning is driven by
 - a. The fastest hardware
 - b. Complete fault prevention
 - c. The cheapest solution
 - d. The requirements of the users
 - e. Absolute security

Part B

1. Network planning allows you to
 - a. Create a perfect network
 - b. Create the largest possible network
 - c. Meet current needs and anticipate maintenance and change
 - d. Guarantee error-free service
2. Describe how network planning can ensure smoother operation later.
3. What are the dangers of running a protocol without management?
4. Name five hardware devices or software tools that are used to gather information for network management and planning

Unit IV – Network Security

What You Will Learn (Objectives)

After completing this unit you will be able to:

- Learn about network security.
- Describe Cryptography, encryption, and authentication
- Know Firewalls, Proxy servers
- Describe Virtual Private Networks

4.1 Introduction

For the first few decades of their existence, computer networks were primarily used by university researchers for sending e-mail and by corporate employees for sharing printers. Under these conditions, security did not get a lot of attention. But now, as millions of ordinary citizens are using networks for banking, shopping, and filing their tax returns, network security is looming on the horizon as a potentially massive problem. In this chapter, we will study network security from several angles, point out numerous pitfalls, and discuss many algorithms and protocols for making networks more secure.

Security is a broad topic and covers a multitude of sins. In its simplest form, it is concerned with making sure that nosy people cannot read, or worse yet, secretly modify messages intended for other recipients. It is concerned with people trying to access remote services that they are not authorized to use. It also deals with ways to tell whether that message purportedly from the IRS saying: Pay by Friday or else is really from the IRS and not from the Mafia. Security also deals with the problems of legitimate messages being captured and replayed, and with people trying to deny that they sent certain messages.

Most security problems are intentionally caused by malicious people trying to gain some benefit, get attention, or to harm someone. A few of the most common perpetrators are listed in figure. It should be clear from this list that making a network secure involves a lot more than just keeping it free of programming errors. It involves outsmarting often intelligent, dedicated, and sometimes well-funded adversaries. It should also be clear that measures that will thwart casual adversaries will have little impact on the serious ones. Police records show that most attacks are not perpetrated by outsiders tapping a phone line but by insiders with a grudge. Consequently, security systems should be designed with this fact in mind.

Network security problems can be divided roughly into four closely intertwined areas: secrecy, authentication, nonrepudiation, and integrity control. Secrecy, also called confidentiality, has to do with keeping information out of the hands of unauthorized users. This is what usually comes to mind when people think about network security. Authentication deals with determining whom you are talking to before revealing sensitive information or entering into a business deal. Nonrepudiation deals with signatures: How do you prove that your customer really placed an electronic order for ten million left-handed doohickeys at 89 cents each when he later claims the price was 69 cents? Or maybe he claims he never placed any order. Finally, how can you be sure that a message

you received was really the one sent and not something that a malicious adversary modified in transit or concocted?

Figure: some people who cause security problems and why

Adversary	Goal
Student	To have fun snooping on people's e-mail
Cracker	To test out someone's security system: steal data
Sales rep	To claim to represent all of Europe, not just Andorra
Businessman	To discover a competitor's strategic marketing plan
Ex-employee	To get revenge for being fired
Accountant	To embezzle money from a company
Stockbroker	To deny a promise made to a customer by e-mail
Con man	To steal credit card numbers for sale
Spy	To learn an enemy's military or industrial secrets
Terrorist	To steal germ warfare secrets

All these issues (secrecy, authentication, no repudiation, and integrity control) occur in traditional systems, too, but with some significant differences. Integrity and secrecy are achieved by using registered mail and locking documents up. Robbing the mail train is harder now than it was in Jesse James' day.

Also, people can usually tell the difference between an original paper document and a photocopy, and it often matters to them. As a test, make a photocopy of a valid check. Try cashing the original check at your bank on Monday. Now try cashing the photocopy of the check on Tuesday. Observe the difference in the bank's behavior. With electronic checks, the original and the copy are indistinguishable. It may take a while for banks to learn how to handle this.

People authenticate other people by recognizing their faces, voices, and handwriting. Proof of signing is handled by signatures on letterhead paper, raised seals, and so on. Tampering can usually be detected by handwriting, ink, and paper experts. None of these options are available electronically. Clearly, other solutions are needed.

Before getting into the solutions themselves, it is worth spending a few moments considering where in the protocol stack network security belongs. There is probably no one single place. Every layer has something to contribute. In the physical layer, wiretapping can be foiled by enclosing transmission lines in sealed tubes containing gas at high pressure. Any attempt to drill into a tube will release some gas, reducing the pressure and triggering an alarm. Some military systems use this technique.

In the data link layer, packets on a point-to-point line can be encrypted as they leave one machine and decrypted as they enter another. All the details can be handled in the data link layer, with higher layers oblivious to what is going on. This solution breaks down when packets have to traverse multiple routers, however, because packets have to be decrypted at each router, leaving them vulnerable to attacks from within the router. Also, it does not allow some sessions to be protected (e.g., those involving on-line purchases by credit card) and others not. Nevertheless, link encryption, as this method is called, can be added to any network easily and is often useful.

In the network layer, firewalls can be installed to keep good packets and bad packets out. IP security also functions in this layer.

In the transport layer, entire connections can be encrypted, end to end, that is, process to process. For maximum security, end-to-end security is required.

Finally, issues such as user authentication and no repudiation can only be handled in the application layer.

Since security does not fit neatly into any layer, it does not fit into any chapter of this book. For this reason, it rates its own chapter.

While this chapter is long, technical, and essential, and it is also quasi-irrelevant for the moment. It is well documented that most security failures at banks, for example, are due to incompetent employees, lax security procedures, or insider fraud, rather than clever criminals tapping phone lines and then decoding encrypted messages. If a person can walk into a random branch of a bank with an ATM slip he found on the street claiming to have forgotten his PIN and get a new one on the spot (in the name of good customer relations), all the cryptography in the world will not prevent abuse. In this respect, Ross Anderson's book is a real eye-opener, as it documents hundreds of examples of security failures in numerous industries, nearly all of them due to what might politely be called sloppy business practices or inattention to security (Anderson, 2001). Nevertheless, we are optimistic that as e-commerce becomes more widespread, companies will eventually debug their operational procedures, eliminating this loophole and bringing the technical aspects of security to center stage again.

Except for physical layer security, nearly all security is based on cryptographic principles. For this reason, we will begin our study of security by examining cryptography in some detail.

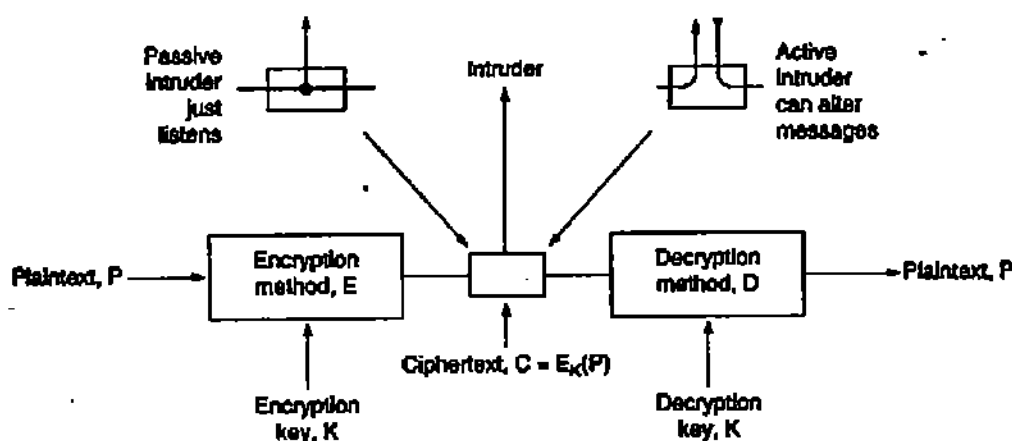
4.2 Cryptography

Cryptography comes from the Greek words for "secret writing." It has a long and colorful history going back thousands of years. In this section we will just sketch some of the highlights, as background information for what follows. For a complete history of cryptography, Kahn's (1995) book is recommended reading. For a comprehensive treatment of the current state-of-the-art in security and cryptographic algorithms, protocols, and applications, see (Kaufman et al., 2002). For a more mathematical approach, see (Stinson, 2002). For a less mathematical approach, see (Burnett and Paine, 2001).

Professionals make a distinction between ciphers and codes. A cipher is a character-for-character or bit-for-bit transformation, without regard to the linguistic structure of the message. In contrast, a code replaces one word with another word or symbol. Codes are not used any more, although they have a glorious history. The most successful code ever devised was used by the U.S. armed forces during World War II in the Pacific. They simply had Navajo Indians talking to each other using specific Navajo words for military terms, for example chay-dagahi-nail-tsaide (literally: tortoise killer) for antitank weapon. The Navajo language is highly tonal, exceedingly complex, and has no written form. And not a single person in Japan knew anything about it.

In September 1945, the San Diego Union described the code by saying "For three years, wherever the Marines landed, the Japanese got an earful of strange gurgling noises interspersed with other sounds resembling the call of a Tibetan monk and the sound of a hot water bottle being emptied." The Japanese never broke the code and many Navajo code talkers were awarded high military honors for extraordinary service and bravery. The fact that the U.S. broke the Japanese code but the Japanese never broke the Navajo code played a crucial role in the American victories in the Pacific.

Figure: the encryption model



The messages to be encrypted, known as the plaintext, are transformed by a function that is parameterized by a key. The output of the encryption process, known as the ciphertext, is then transmitted, often by messenger or radio. We assume that the enemy, or intruder, hears and accurately copies down the complete ciphertext. However, unlike the intended recipient, he does not know what the decryption key is and so cannot decrypt the ciphertext easily. Sometimes the intruder can not only listen to the communication channel (passive intruder) but can also record messages and play them back later, inject his own messages, or modify legitimate messages before they get to the receiver (active intruder). The art of breaking ciphers, called cryptanalysis, and the art devising them (cryptography) is collectively known as cryptology.

It will often be useful to have a notation for relating plaintext, ciphertext, and keys. We will use $C = E_K(P)$ to mean that the encryption of the plaintext P using key K gives the

ciphertext C . Similarly, $P = D_K(C)$ represents the decryption of C to get the plaintext again. It then follows that

$$D_K(E_K(P)) = P$$

This notation suggests that E and D are just mathematical functions, which they are. The only tricky part is that both are functions of two parameters, and we have written one of the parameters (the key) as a subscript, rather than as an argument, to distinguish it from the message.

A fundamental rule of cryptography is that one must assume that the cryptanalyst knows the methods used for encryption and decryption. In other words, the cryptanalyst knows how the encryption method, E , and decryption, D , of figure work in detail. The amount of effort necessary to invent, test, and install a new algorithm every time the old method is compromised (or thought to be compromised) has always made it impractical to keep the encryption algorithm secret. Thinking it is secret when it is not does more harm than good.

This is where the key enters. The key consists of a (relatively) short string that selects one of many potential encryptions. In contrast to the general method, which may only be changed every few years, the key can be changed as often as required. Thus, our basic model is a stable and publicly-known general method parameterized by a secret and easily changed key. The idea that the cryptanalyst knows the algorithms and that the secrecy lies exclusively in the keys is called Kerckhoff's principle, named after the Flemish military cryptographer Auguste Kerckhoff who first stated it in 1883 (Kerckhoff, 1883). Thus, we have:

Kerckhoff's principle: All algorithms must be public; only the keys are secret

The nonsecrecy of the algorithm cannot be emphasized enough. Trying to keep the algorithm secret, known in the trade as security by obscurity, never works. Also, by publicizing the algorithm, the cryptographer gets free consulting from a large number of academic cryptologists eager to break the system so they can publish papers demonstrating how smart they are. If many experts have tried to break the algorithm for 5 years after its publication and no one has succeeded, it is probably pretty solid.

Since the real secrecy is in the key, its length is a major design issue. Consider a simple combination lock. The general principle is that you enter digits in sequence. Everyone knows this, but the key is secret. A key length of two digits means that there are 100 possibilities. A key length of three digits means 1000 possibilities, and a key length of six digits means a million. The longer the key, the higher the work factor the cryptanalyst has to deal with. The work factor for breaking the system by exhaustive search of the key space is exponential in the key length. Secrecy comes from having a strong (but public) algorithm and a long key. To prevent your kid brother from reading your e-mail, 64-bit keys will do. For routine commercial use, at least 128 bits should be used. To keep major governments at bay, keys of at least 256 bits, preferably more, are needed.

From the cryptanalyst's point of view, the cryptanalysis problem has three principal variations. When he has a quantity of ciphertext and no plaintext, he is confronted with the ciphertext-only problem. The cryptograms that appear in the puzzle section of newspapers pose this kind of problem. When the cryptanalyst has some matched ciphertext and plaintext, the problem is called the known plaintext problem. Finally, when the cryptanalyst has the ability to encrypt pieces of plaintext of his own choosing, we have the chosen plaintext problem. Newspaper cryptograms could be broken trivially if the cryptanalyst were allowed to ask such questions as: What is the encryption of ABCDEFGHIJKL?

Novices in the cryptography business often assume that if a cipher can withstand a ciphertext-only attack, it is secure. This assumption is very naive. In many cases the cryptanalyst can make a good guess at parts of the plaintext. For example, the first thing many computers say when you call them up is login: . Equipped with some matched plaintext-ciphertext pairs, the cryptanalyst's job becomes much easier. To achieve security, the cryptographer should be conservative and make sure that the system is unbreakable even if his opponent can encrypt arbitrary amounts of chosen plaintext.

Encryption methods have historically been divided into two categories: substitution ciphers and transposition ciphers. We will now deal with each of these briefly as background information for modern cryptography.

4.3 Encryption

Encryption keeps people from reading your data unless you want them to. When you encrypt something, you translate it into a secret code, called cipher text. In order to decipher it, the person you are sending it to needs a *key*. Only someone with the key can read your data.

Basic Encryption

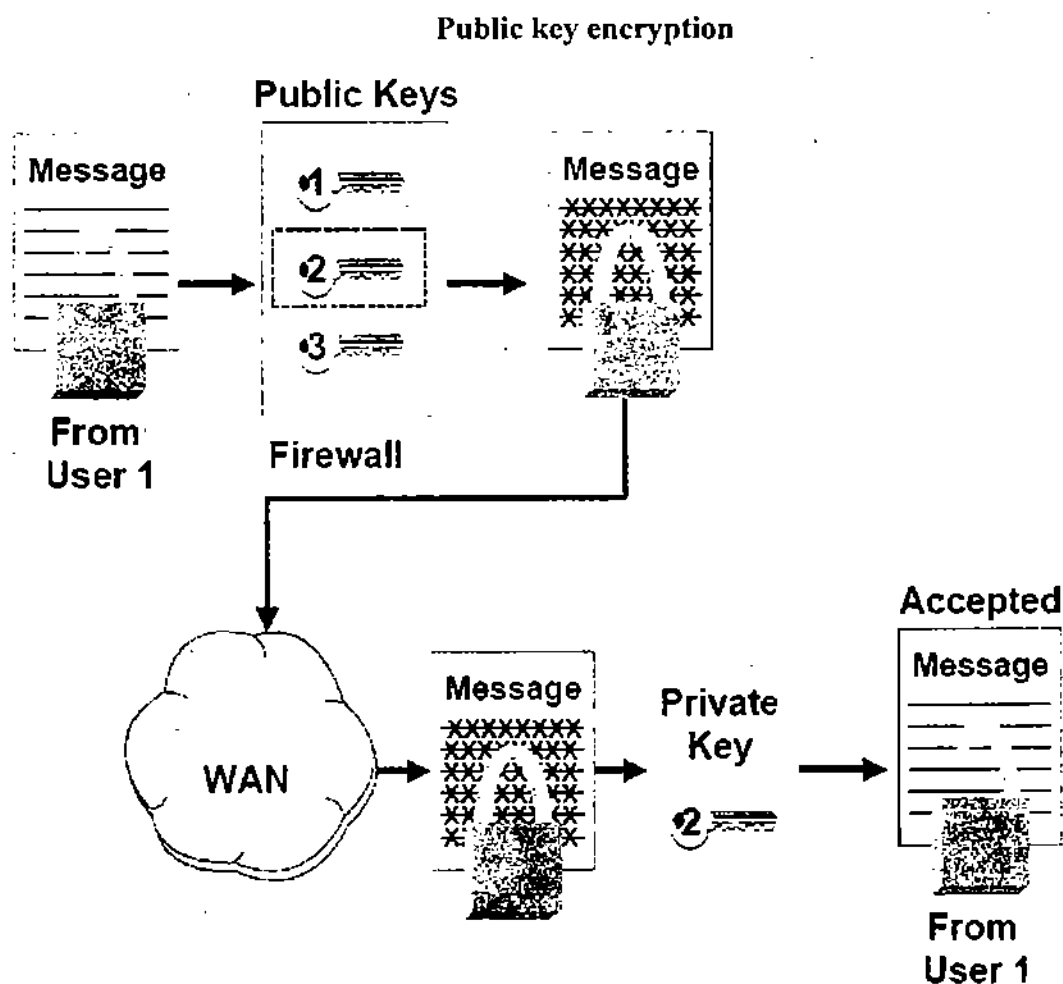
Plain Text	Encryption Key	Cipher Text	Cipher Text	Encryption Key	Plain Text
R	$=18+5=23=$	W	W	$=23-5=18=$	R
E	$=5+5=10=$	J	J	$=10-5=5=$	E
A	$=1+5=6=$	F	F	$=6-5=1=$	A
D	$=4+5=9=$	I	I	$=9-5=4=$	D
T	$=20+5=25=$	Y	Y	$=25-5=20=$	T
H	$=8+5=13=$	M	M	$=13-5=8=$	H
I	$=9+5=14=$	N	N	$=14-5=9=$	I
S	$=19+5=24=$	X	X	$=24-5=19=$	S

You might have used secret codes when you were younger (or maybe you still do) that worked like this: You translate a message into code using a key. The key might include several steps such as:

1. Change the letters to numbers, 1 to 26.
2. Add five to each number, for numbers larger than 26 start back at 1.
3. Convert the number back to a letter.

You whisper the key to your friend so no one else can hear, or you pass it to him/her in a note that s/he swallows after memorizing. Then you send the message. This is called secret key or symmetric encryption. Both the sender and the receiver know the key but no one else does. Secret key encryption works pretty well when you can whisper the key to the receiver, but it doesn't work so well when you want to give the key to someone far away across a WAN. If you could send them the key without worrying about it being intercepted then you probably don't need encryption to begin with.

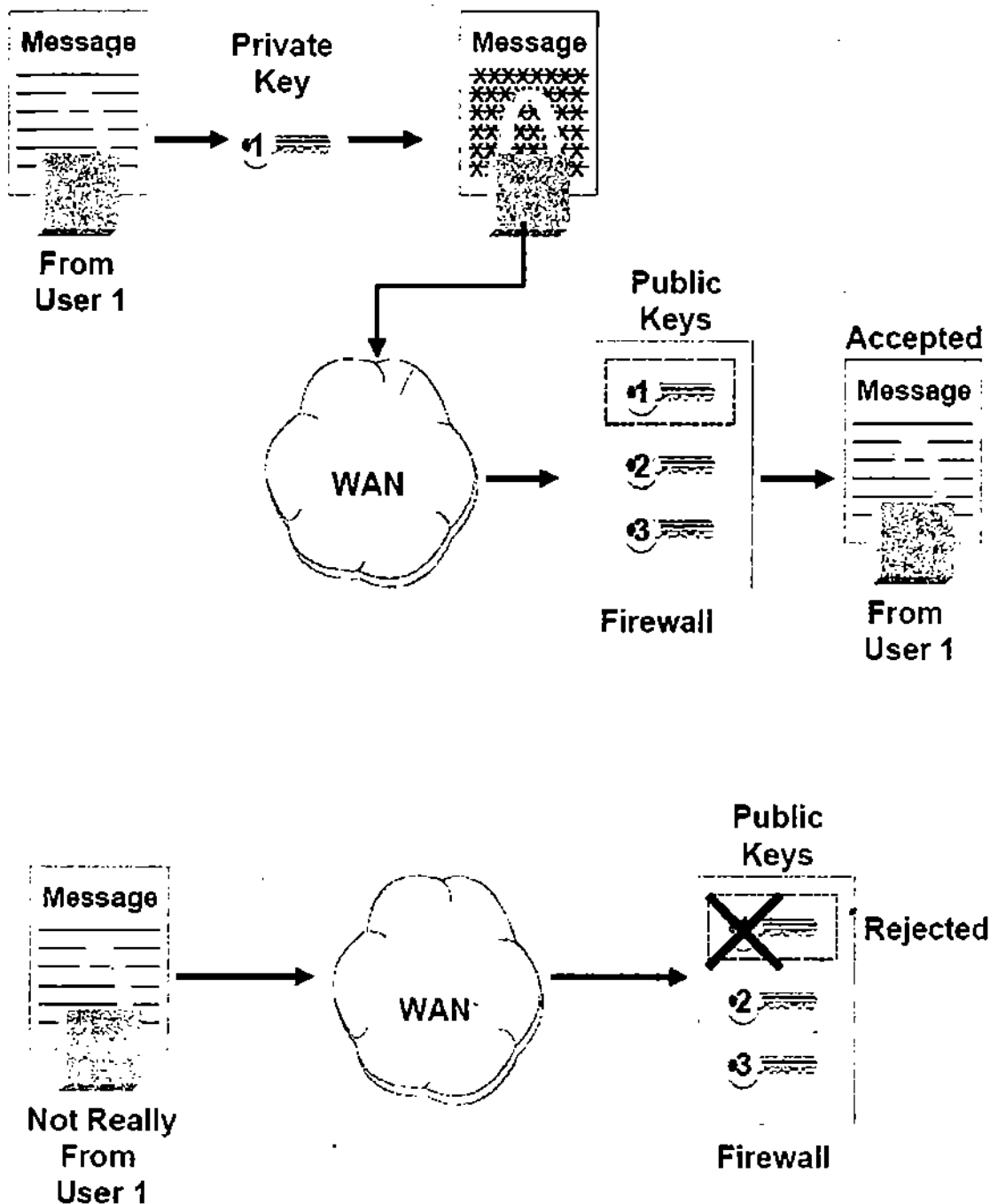
A safer method of encryption is called public key encryption. With public key encryption, each user has two keys, one public key that everyone knows and one private key that only the sender knows. Each key is a very large number that is used in a function to change the characters in the message. When someone sends you a message enciphered with your public key, only you can decipher it using your private key.



4.4 Authentication

Authentication is a way of making sure that information sent from a person actually comes from that person. A malicious person may want to put information onto a LAN that doesn't belong there, for example, a virus. To gain access to the LAN, the person might pretend to be an authorized user by using the authorized user's address. To an unprotected LAN, that information will look like it came from the authorized user. To prevent this kind of intrusion from happening, a LAN might use *data authentication* or *user authentication*.

Data Authentication



Public key encryption, described above, also works in reverse for data authentication. The firewall using public key encryption can be set to accept only encrypted information from certain addresses. The encrypted messages are deciphered using public keys. In order for a message to be decipherable using a public key, it must have been encrypted using the matching private key, which means it is authentic.

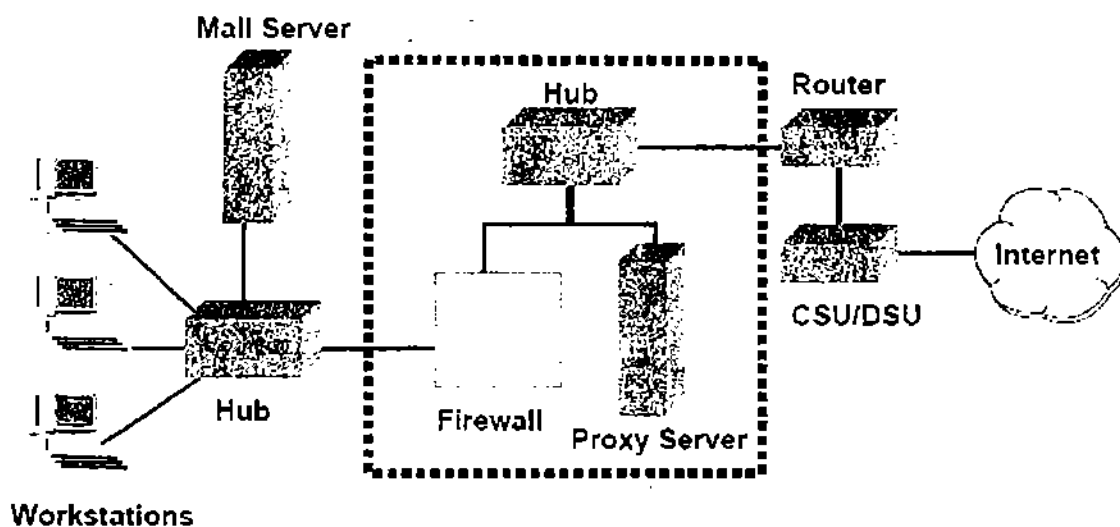
Another form of authentication prevents unauthorized users from logging onto a network. The most common type of *user authentication* is a user password that must be typed before gaining access.

4.5 Firewalls

A firewall protects the security of a network by restricting the information that is allowed to pass between the LAN and the WAN. The simplest firewall is a router that sits between the WAN and the LAN. All data from the WAN must pass through the router. The router may perform encryption and authentication. It may also be set to only allow transmissions that:

- Are coming in or going out.
- Are of a certain type (email, telnet, ftp, and so on).
- Contain certain data or *do not* contain certain data.
- Come from a particular address or *do not* come from a particular address.
- Arrive before or after a certain time of day.

A DMZ

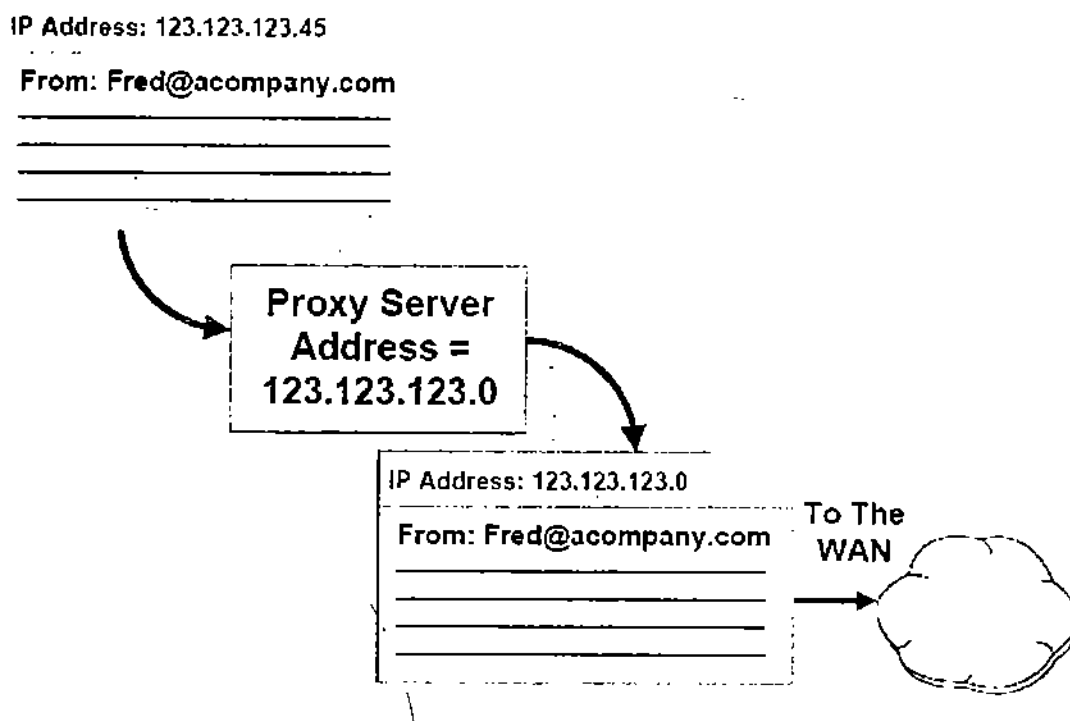


Some companies may want stronger protection. They might hire a company that specializes in security to set up a small network, called a "DMZ", or Demilitarized Zone. Any information coming in from the WAN must pass through the DMZ before it gets onto the LAN. The DMZ contains a firewall server that does authentication and encryption using special software, as well as a proxy server.

4.6 Proxy Servers

Each computer on a network connected to the Internet must have an address. If a *hacker* learns the address of a particular computer inside your LAN, that person can use that information to get inside the LAN. When you send email, your messages contain the address of your computer. When you browse the web, every click of your mouse is actually a message sent across the Internet to the server that contains that web page. The mouse click also includes your computer's address. Companies that send junk email, called *spam*, collect the IP addresses of computers that have accessed certain websites. Then they send email directly to those addresses.

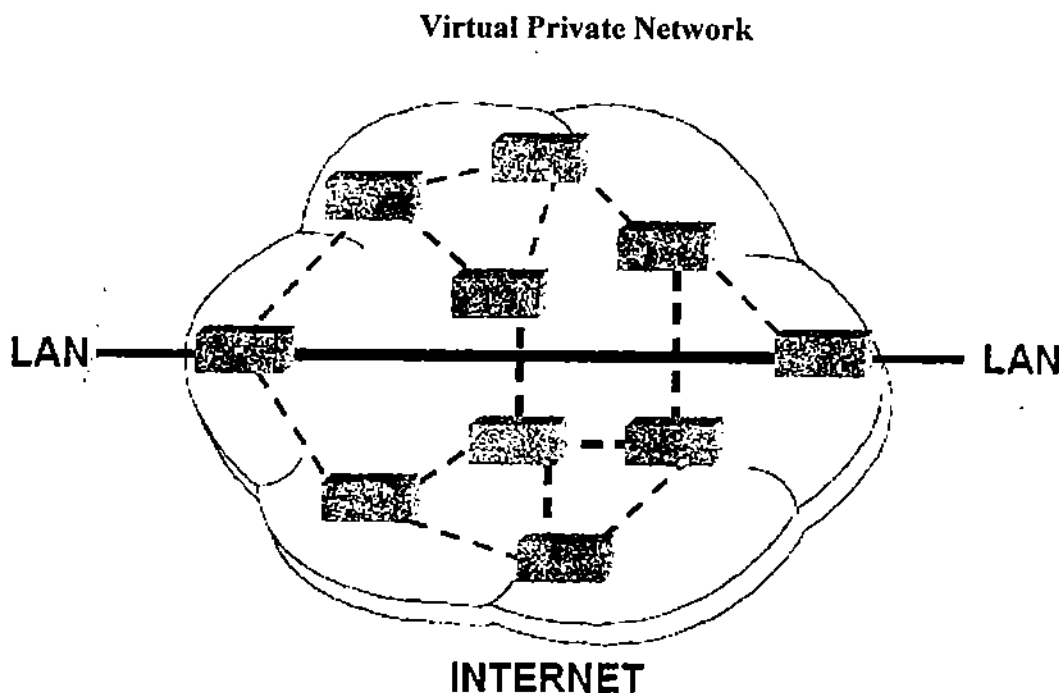
Proxy Server



A proxy server hides the address of the sending computer before passing data out to the Internet. It does this by using its own address to replace the address of any computer on the LAN. The replacement address is called a proxy address. When a *spammer* tries to use the address, the mail goes to the proxy server instead of directly to a workstation on the LAN. The proxy server can then be set to reject messages from addresses that are known to be sources of spam.

4.7 Virtual Private Networks (VPNs)

In the WAN Configurations lesson, you learned that some WANs called Virtual Private Networks (VPNs) use the Internet. Even though the data travels across the Internet, to a user at either end, a VPN seems like a leased line: the connection is always open and the data always goes to the same place. Sending data across the Internet in a VPN is known as "tunneling." By using a tunnel, packets from specific sources, for example, a credit card approval from a bank, can be automatically separated from all the other data travelling on the network. The tunneled data can also be given higher quality of service.



Internet Service Providers (ISPs) offer VPNs. Since a LAN connected to the Internet can communicate with any other LAN connected to the Internet, a VPN is actually a way of making that connection more secure. To make the VPN private, the data is encrypted and authenticated.

Preparing data to be sent through a VPN is typically the job of the router that runs a tunneling protocol. A tunneling protocol puts data packets from a LAN into packets that transmit over the Internet. A router isn't necessary, even laptops can run the software necessary to use a VPN. Windows 95, 98, and NT include this software.

4.8 Summary

This unit introduces the concept of network security and gives a detailed look at network security aspects. Next, it describes Cryptography, Encryption, and Authentication. E-mail security can be achieved by a combination of the techniques. The cryptographic tools can be used to secure network traffic. Then we learn about firewalls and proxy servers. Finally, virtual private networks (VPNs) are briefly explained.

4.9 Review Questions

1. Why is public key encryption safer than basic encryption?
2. What is the purpose of authentication?
3. What is the purpose of firewall?
4. What do you understand by network security?
5. Describe cryptography.
6. What are proxy servers?
7. Describe virtual private networks (VPNs).
8. List the people who cause security problems. Why they do so?

Objective Questions

1. To verify that the sender of an encrypted message is who they say they are
 - a. Use the sender's public key to decrypt the message
 - b. Use the sender's private key to decrypt the message
 - c. Use your public key to decrypt the message
 - d. Use your private key to decrypt the message
2. To send an encrypted message that only the recipient can read
 - a. Use the recipient's public key to encrypt the message
 - b. Use the recipient's private key to encrypt the message
 - c. Use your public key to encrypt the message
 - d. Use your private key to encrypt the message
3. Which function is not provided by a firewall?
 - a. Encryption and authentication
 - b. Blocking connections to specific addresses
 - c. Blocking connections from specific addresses

- d. Hiding the addresses of local computers when they connect to other networks
4. Which function does a proxy server provide?
- a. Encryption and authentication
 - b. Blocking connections to specific addresses
 - c. Blocking connections to specific addresses
 - d. Hiding the addresses of local computers when they connect to other

NOTES

NOTES



Uttar Pradesh
Rajarshi Tandon Open University

BLISS-03

**Computer Network and
Security Maintenance**

BLOCK

5

Software Solution of Peripherals

Unit 1

Overviews of System Software

5

Unit 2

Software Solution of Devices

19

Unit 3

Basic Input Output System Handling

28

Unit 4

Installing and Configuring the Device Driver

39

BLOCK INTRODUCTION

Day to day, advancements in the computer field are going to improve. Behind all these figure, only software is responsible to put the application of computer in different field. This block emphasizes the fundamentals of software and its types. We have introduced little attention about software, system software and application software. Finally, we have also discussed about the roles of BASIC INPUT OUTPUT SYSTEM (BIOS) in maintenance of computer. When the students are interested to learn maintenance of computer, thus should know about basics of system software.

System Software and Operating System Concept: Concept of system software is more necessary, when candidates have moved towards computer field.

Software Concept: Concept of software is also necessary to deal maintenance of computer properly.

Device Driver Concept: Concept of device drivers is to be introduced for interfacing hardware with software properly.

BIOS handling: It is also necessary to discuss the Basic Input Output System (BIOS), which is capable to control I/O device of computer.

Since, this block is introduced to discuss only software orientation with hardware, so we will focus on the basics of software as well as system software.

The first unit of this block is concentrated on basics of system software, which create conceptual view of system software and also put the comparative picture of system software and application software.

In the second unit of this block, we have mainly given the conceptual view software orientation towards devices.

In the third unit of this block, we have devoted on what is a device driver and how it is useful to maintain synchronization with hardware of computer.

In the fourth unit of this block, we have discussed about Basic Input Output System (BIOS), and how it is responsible to maintain coordination in between hardware and software.

Suggestion for further reading

1. Foundation of Information Technology, Second Edition, D.S. Yadav
2. Information Technology, Third Edition, Atul Kahate
3. PC Upgrade & Maintenance Guide, Fifteen Edition, Mark Minasi
4. Modern All About Motherboard, Manhar Lotia, Pradeep Nair

UNIT-1: OVERVIEWS OF SYSTEM SOFTWARE

Structure:

- 1.0 Objectives
- 1.1 Introduction
- 1.2 System Software
 - 1.2.1 System Control Software
 - 1.2.1.1 Operating System Environments
 - 1.2.1.2 Multiprogramming Environments
 - 1.2.1.3 Time-Sharing Environments
 - 1.2.1.4 Multiprocessing Environments
 - 1.2.2.5 Operating System and its Functions:
 - 1.2.2.5.1 MS-DOS
 - 1.2.2.5.2 WINDOWS
 - 1.2.2.5.3 Windows 95
 - 1.2.2.5.4 Windows NT
 - 1.2.2.5.5 UNIX
 - 1.2.2.5.6 OS/2
- 1.3.2 System Support Software
- 1.3.3 Systems Development Software
- 1.4 Summary
- 1.5 Model Question

1.0 Objectives

At end of this unit candidate will be able to:

- Understand about the basic concept of system software.
- Distinguish in between system software and application software.
- Explain the various kind of system software.
- Know the concept of various types of processing modes.
- Analyze statistics of various operating systems.
- Generate awareness about system support software.
- Identify conceptual view of system development software.

1.1 INTRODUCTION

Computers are playing an important role in our life today and their importance is increasing day by day. They are used in finding solution to a significant range of problems in all spheres. In the solution of familiar problems computer are saving considerable time and money.

In most organizations, the computer is a very valuable resource. Among the resources that a computer has processing time, storage space, printers and terminals. Largely, a type of system software called an operating system performs the management of these resources. When users interact with the computer, most of this interaction has been done with the help of system software. System software directly affects the ease with which users interact with a computer. Furthermore, system software is a very significant determinant of whether or not a particular computer's resources are used optimally. For these reasons, it is important that users of computer systems have a basic knowledge of system software.

Here, we will first distinguish between system software and application software then we will explore the functions, components, and types of operating systems. Finally, we will see how operating systems make possible multiprogramming, virtual storage, timesharing, multiprocessing, multitasking and multi-user computer systems.

Software can be classified into two broad categories: system software and application software. System software performs computer-related tasks, such as managing input and output devices; application software performs people-related tasks, such as human resources and marketing. A more detailed discussion of these two types of software follows.

1.2 System Software

System software is a set of programs that manage the resources of a computer system (processing time, storage space, and so on) so that they are used in an optimal fashion, provide routine services such as copying data from one file to another, and assist in the development of application programs. System software consists of general programs that assist the computer in the efficient control, support, development and execution of application programs.

System software began to be used extensively with the second-generation computers in the early 1960s. In the primary stage, human operators controlled the operation of a computer. These operators monitored the processing of each job. Typically, when a job ended, a bell rang or a light flashed to indicate that another job should be input to the computer and started by the operator. If a job ended while the operator was having a tea or coffee break, the computer might have sat idle for five or ten minutes or longer. In addition, the operator had to activate each peripheral device when the computer needed that device. This type of human intervention wasted large amounts of computer time and human resources. To automate these functions, companies developed programs called operating systems. These programs are stored partially in primary storage and indirect access secondary storage devices (DASD) so the computer can access them immediately when they have needed. With operating systems, a queue of jobs that are awaiting execution can read onto a disk. The operating system will start each job when system resources are available for its execution. Since human intervention has eliminated, computer idle time has been significantly reduced.

System software performs the basic functions necessary to start and operate a computer. It controls and monitors the various activities and resources of a computer and makes it easier and more efficient to use the computer. System software can be classified into three categories.

- **System Control Software:** System control programs control the execution of programs, manage the storage and processing resources of the computer, and perform file management and monitoring the functions. The most important of these programs is the operating system; other examples are database management systems (DBMS) and communication monitoring software.
- **System Support Software:** System support programs provide routine service functions to the other computer programs and computer users. Examples are utilities, librarians, performance of monitors and job accounting.
- **System Development Software:** System development programs assist in the creation of application programs. Examples are language translators such as BASIC interpreter, language compilers and application generators such as IBM's Application Development Facility.

1.2.1 System Control Software:

System control software includes programs that monitor, control, coordinate, and manage the resources and functions of a computer system. The most important system control software is the operating system.

A very important part of the system control software is the operating system, which performs many functions. Two of its more critical tasks are starting the computer (initial program load or bootstrap) and initiating the processing of each job. To understand these functions" let us examine the Microsoft Disk Operating System (MS-DOS) used in the IBM PC. ROM-BIOS provides very fundamental services needed by the computer such as self-test of memory, starting up (booting) the computer and input/output services between the central processing unit and peripheral devices such as printers and disks. It is stored permanently in ROM. The diskette boot record is a very short and simple program stored at the beginning of the DOS diskette. Its purpose is to begin the process of loading the operating system when the PC is first turned on, non-resident means that is not stored in RAM while the PC is operating. Both IBMBIO.COM and

IBMDOS.COM are extensions of ROM-BIOS. They provide additional input/output interfaces with peripheral devices. They are stored on the DOS diskette, but they are hidden files. A hidden file has not displayed when a DIR (directory) of the diskette is produced. They are resident in RAM while the PC is operating. The primary job of COMMAND.COM is to process and interpret the commands that you type into DOS. It also contains the programs that execute several DOS commands. These programs are resident in RAM. The final part of DOS, the external commands are not resident in RAM these are moved from the DOS diskette to RAM whenever they are needed.

In the following figure, steps of PC-DOS or MS-DOS are shown.

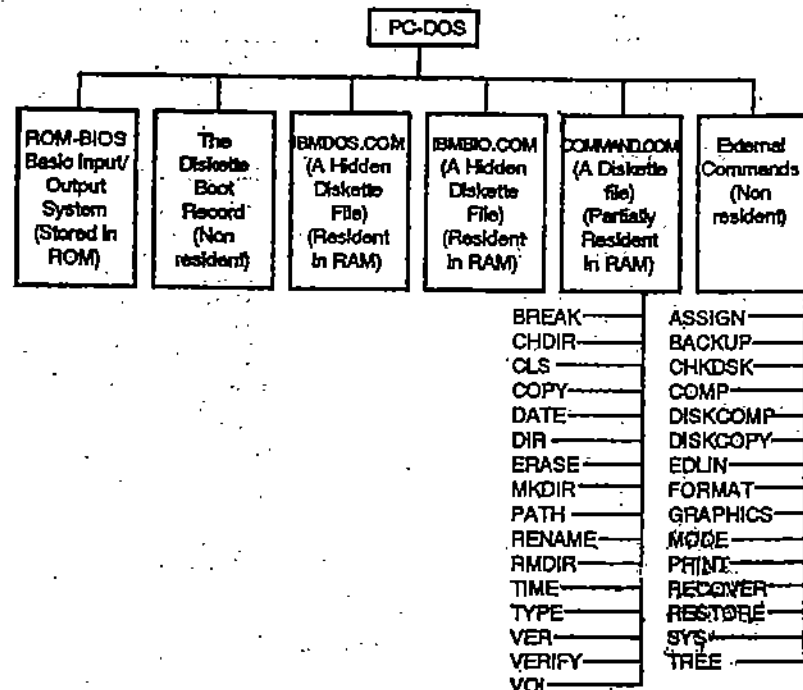


Figure : MS-DOS Operating System

Steps of MS-DOS are also described as below.

- Bootstrap loader resident in ROM loads operating system.
- Operating system loads file directory.
- Operating system indicates on monitor that computer is in system mode.
- Operating system reads file directory to check availability and size of available software.
- Operating system checks to see if space is available in primary memory (yes).
- Operating system causes software to be loaded.
- Computer is in software mode and that fact has displayed on monitor.
- Software decodes instruction.
- Instruction requires access to disk software requests assistance.
- Operating system checks availability of file and of space in primary memory.
- File is loaded into primary memory.
- Control has restored the software and fact has displayed on monitor.

The operating system transfers a file directory from disk memory to primary memory in the file directory is listed the address, or position, of every program and data

file recorded on the disk. In response to the next instruction, the operating system finds the software on the disk and after making certain there is enough space for it, loads into primary memory the user is notified that the software is ready. (Some personal computers perform automatically as part of the switching-on sequence.) The operating system is called on to load the application program itself. Now, with the software again in control, the application program can be run. Output will be a new data file in primary memory, which can be transferred to disk storage.

Operating systems on minicomputers and mainframe computers also perform job-scheduling functions by examining the priority of each job awaiting execution. Jobs with higher priorities are executed first. Access security is also a very important function of the operating system. The function is carried out through various password schemes that identify valid users and determine which data files, they may require to access.

The storage of a mainframe operating system illustrates now. The system residence device stores the complete operating system. Today this is usually a disk unit as portions of the operating system are needed for execution; they can Core-Image Library Operating System Programs: Initial Program Loader Supervisor Job-Control Program input/Output Control Program.

Storage of a mainframe operating system is loaded readily into primary storage. (A core-image library means that the programs stored on the disks can be moved directly into primary storage without modification. This improves the speed of the operating system can be moved into the supervisor transient area as needed).

The following figure has shown the storage capacity of mainframe computer. It has **very** large capacity.

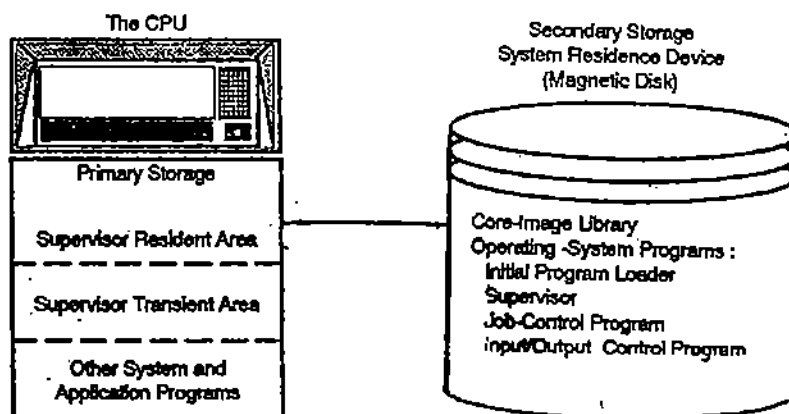


Figure: Shown the storage of Mainframe Operating System

There are generally four types of operating system programs: the Initial Program Loader (IPL) the supervisor the job-control program, and the input/output control program. In mainframe computers, the bootstrap program is known as the initial program loader. Its purpose is to start up operations. It performs this function by reading the resident portion of the supervisor from secondary storage and loading it into primary storage. Since the operating system is constantly supervising and monitoring the computer, frequently used portion of the operating system (called the resident supervisor) is stored in primary storage while the computer is operating. Other parts of the supervisor are less frequently used. They are stored only temporarily in the supervisor transient area of primary storage, when they are in use.

Once the resident portion of the supervisor is loaded into primary storage, control is passed to the supervisor and the operation of the computer begins. The supervisor programs (often called monitor or executive programs) are the principal managers in an operating system. They organize and control the flow of work by initiating and controlling the execution of other computer programs.

As operating system's software replaced human operators in the control of mainframe computers, new languages were developed to enable users and programs to communicate with the operating system. A job-control language (JCL) requires that the user include several job-control statements along with a program. The statements identify the job and its steps to specify the system resources to be used (for example, expected run time, input/output device to be used, and memory space required). Job-control language also describes the data sets or files that are to be used in the various job steps.

Concept Operating System and its Functions: An operating system is a collection of programs used to control a computer system. Part of the operating system is stored in secondary memory on a floppy disk or hard disk; while part is stored permanently in read only memory (ROM). When the computer is first started, it automatically begins to execute a program contained in the operating system portion of ROM. This program loads part of the operating system into random access memory (RAM) and the computer execute this part of the operating system. The operating system program now controls the computer.

In other word, we may also define an operating system is a complex set of software modules that manages the overall operations of a computer. A master control program acts as a manager, a housekeeper and a traffic controller for the computer system. Depending on the type of computer (mainframe, personal computer and so on), the operating system performs a number of functions, such as allocating resources for running different software programs, tracking the use of different computer resources, ensuring optimal utilization of various resources and acting as an interface between the user and the computer. Application programs, such as word processors and spreadsheets run under the supervision of the operating system. In the following figure, we can see how the hardware, system software and application software are associated.

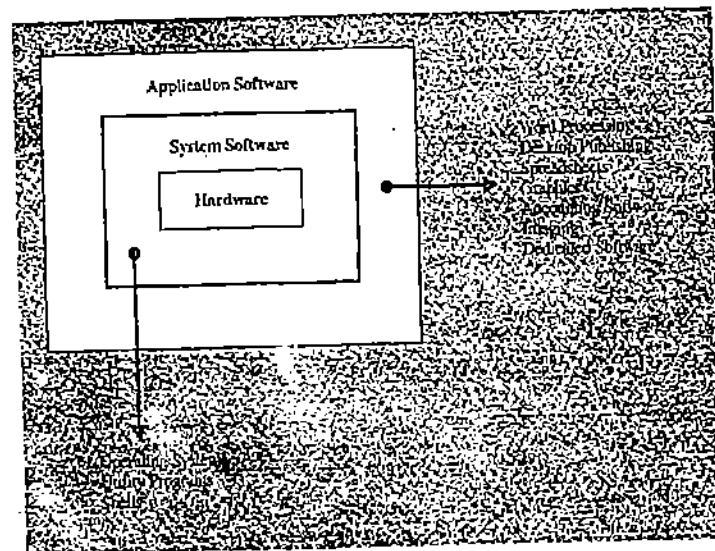


Figure: The relationships among hardware, system software, and application software

An operating system loads programs, performs and manages input/output operations, manages files, monitors and manages the use of computer memory, allocates resources for various computer functions, and resolves conflicts. Some of these functions are carried out only by organization-wise systems and workgroup systems, such as mainframes and minicomputers; others are performed by personal information systems, such as PCs. Let us look at some of these functions.

The loaded program of operating system is executed in primary memory and sends a message to the user when the execution is completed. It also notifies the user of any errors in the system that it encounters during program execution.

The operating system allocates and manages input/output devices and provides their instructions, such as "start printer" and "rewind tape", to various programs. The operating system provides access to different I/O devices and "releases" these devices when a task is completed so that they can be used by other programs. If user intervention is necessary because of a loose printer cable connection or lack of paper in the printer, it sends a message to the user. Another function of an operating system is to manage files (a file is a place in computer memory where data and instructions are stored). The computer reads, creates, deletes, merges, and renames files and performs other file, related tasks with help from the operating system.

The operating system also manages hard disk storage so that users can create, execute, save, and retrieve various applications. Hard disk storage management is an important consideration in ensuring the efficient utilization of computer memory managing computer resources such as the CPU, primary memory, secondary storage, I/O devices and other peripherals is another function of the operating system as the resources in a computer system must be carefully synchronized. Operating systems for organization wise systems and work-group systems track the use of computer resources for cost accounting and statistical purposes. Statistics on resource use can help managers to better plan, implement and coordinate strategies that will lead to full utilization of existing resources. For example, if usage statistics show that computer memory has over utilized while the CPU has under utilized, a manager can improve operational efficiency by expanding computer memory.

- Non-multi Programming
- Multi Programming

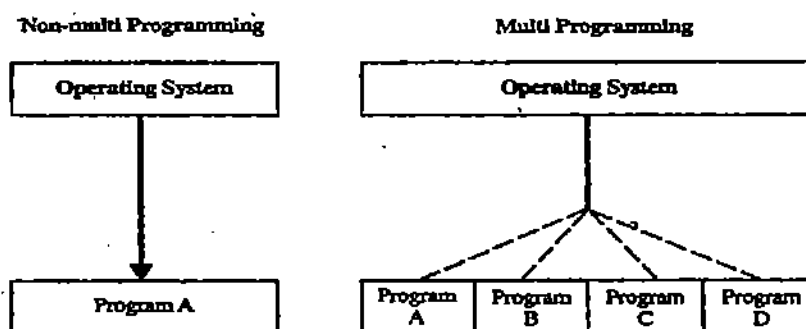


Figure : Concept Multi and Non-multiprogramming

The above figure shows the concept of multiprogramming and non-multiprogramming.

Finally, the conflict resolution is an important function of an operating system for mainframes. When resolving a resource conflict, the operating system serves as an arbitrator by taking into account factors such as the criticality of the application and the priority of the user, and allocates resources accordingly. It ensures that limited system resources, such as the CPU and the I/O devices, have been allocated among different programs according to predetermined criteria. For example, a payroll program often gets higher priority than a program that calculates employee travel expenses; the chief financial officer may get higher priority than a part time worker may. Summary of few important functions of the operating system is given in the following table.

Function	Description
Load Programs	Moves programs has executed into primary memory and indicate when program execution is complete.
Manage I/O operations	Enables and manages the use of multiple I/O devices. Ensure the efficient utilization of I/O devices.
Manage files	Allows application programs to read and write files. Performs other file related operations such as copying, deleting, moving and so on.
Manage memory	Allocate memory so users can create, execute, and retrieve different applications. Ensures that memory is fully and efficiently utilized.
Detect errors	Bring system and execution errors to the attention of the user so that timely action has taken.
Allocate resources	Manage all the computer resources, including the CPU, primary memory, secondary storage, I/O devices and other peripherals.
Monitor resource use	Tracks use of computer resources for cost accounting and statistical purposes. Mostly have done for enterprise wide systems.
Resolve conflicts	Act as an arbitrator among competing programs, taking into account factors such as the criticality of the application or the priority of the program.

Table: A summary of the some important functions of an operating system

1.2.1.1 Operating System Environments:

1.2.1.2 Multiprogramming Environments: Multiprogramming emerged in an effort to address the wide discrepancy between the processing speed of the CPU and input/output devices. The CPU is many times faster than input and output devices. Hence, while input/output operations had been executed the CPU sits idle. This is a tremendous waste of computer resources and multiprogramming helped to overcome this bottleneck and increase the efficiency of computer operations.

A multiprogramming environment works in the following way: A number of programs are executed and kept in primary storage. The operating system selects and begins execution of the first instruction or element in the first program. When, it encounters an I/O instruction, the operating system hands off the instruction to the I/O device. While the I/O device is processing the instruction, the CPU selects the next program from the program queue and begins execution of the second program. When the I/O processing of the first program is complete, the operating system notifies the CPU, which then switches back to the first program and continues execution until another I/O activity has encountered. This greatly increases the utilization of the CPU and ensures that the slow processing speed of I/O devices does not affect CPU utilization. Therefore, in a

multiprogramming environment the CPU switches between multiple programs, although at any given time the CPU is executing only one program.

1.2.1.3 Time-Sharing Environment: Time-Sharing Environment is similar to multiprogramming environment. In this environment, different users get a slice of the CPU's time. In time-sharing, like multiprogramming, the CPU is not dedicated to one user but look after the many users. However, the CPU is so fast that users are unaware that the CPU is switching between different users because of the tremendous speed with which the CPU processes instructions. Time-sharing systems had been developed so that users could interface in an interactive or conversational mode directly with the CPU through dumb terminals. When the user types in a processing request, the CPU processes the request and sends a response, where ever necessary, to the user's terminal. Time-sharing greatly facilitates testing code and developing system applications because users could locate and correct errors in a matter of a few minutes, instead of having to wait for hours to get some of CPU time.

1.2.1.4 Multiprocessing Environments: A multiprocessing system has several processors or CPUs and they work simultaneously. Multiprocessing is ideally suited for complex and computationally intensive operations that require a large amount of processing. Note that multiprogramming and multiprocessing are not mutually exclusive.

In the following table, brief descriptions about multiprocessing and multiprogramming are given.

Operating System	Environment Description
Multiprogramming	Multiple users can run multiple programs on a single-CPU computer at the same time. The CPU switches between programs; however, at any given time it is executing only one program.
Multiprocessing	A multiprocessing system has a number of processors that process data and instructions; unlike systems that have only one CPU. Ideally suited for complex and computationally intensive operations that require extensive processing

Table: A brief description of Multiprocessing and Multiprogramming

Today there are many types of operating systems in the market. Some have been designed for microcomputers and workstations; others have targeted for mainframes and minicomputers. In the next section, we look at some popular operating systems. Examples of Operating Systems: MS-DOS, Windows, UNIX and OS/2 are the most popular operating systems for PCs. While some of these operating systems, such as MS-DOS, are exclusively for PCs, other operating systems, such as UNIX, are available for both large and small computers. Following examples have considered understanding of operating systems.

1.2.2.5.1 MS-DOS:

MS-DOS stands for Micro Soft Disk Operating System. It is a popular operating system for machines based on Intel processor chips. It is a single-user system. In this, only one user at a time can use the computer. DOS has been in use for

many years and the Windows operating system (covered below) has been built on DOS. The commands in DOS are somewhat cryptic and not very user-friendly. Two important files in DOS are CONFIG.SYS and AUTOEXEC.BAT. A CONFIG.SYS file contains commands that setup various parameters for your system. Each time you start your system, DOS searches for CONFIG.SYS file; if it finds one, it uses the commands there to configure your system. After searching for the CONFIG.SYS file, DOS searches for an AUTOEXEC.BAT file. If DOS finds an AUTOEXEC.BAT file, it uses the commands to further define the way it starts your system. In other words, the AUTOEXEC.BAT file sets certain parameters that facilitate using the computer. If an AUTOEXEC.BAT file is not included, then the user will be in the default mode when he or she starts the computer. For example, if a user wants to open a word processing application each time he starts the computer the command for this would be included in the AUTOEXEC.BAT file.

1.2.2.5.2 WINDOWS:

The Windows operating system, developed by Microsoft, expands on the DOS operating system; users can activate programs from Windows using icons (or symbols). An icon is a picture on the screen that represents an action or application that the computer can implement. Windows is a graphical user interface (GUI) that uses the point-and-click method (i.e. the use of a mouse to point at an icon or menu item and click on it) to execute different commands, such as the file commands open, close, delete and move.

In Windows, each application appears in its own window. For example, word processing can appear in one window, a spreadsheet in another window, and a graphics program in a third. A user can easily move between windows to switch applications. Further, more Windows is a highly integrated environment, in which different applications have the same "look and feel", so users familiar with one application can easily work in other applications. For example, in Windows the symbol for closing a file or a document is the same in a word-processing application or in a spreadsheet application. Many companies now use Windows as their operating system. For example, Boston Chicken uses Windows-based PCs for applications such as customized forecasting, scheduling and inventory management.

1.2.2.5.3 Windows 95:

The Windows Version 4.0 is in many ways similar to Windows. Yet it is a radical departure from Windows in that it is independent of DOS and allows for "plug-and-play", unlike Windows. Like Windows, Windows 95 represents programs with icons; when a user clicks on an icon, the system recognizes and opens the application associated with it. For example, suppose you have filed a called RESUME.DOC in Windows 95. If you simply click on the file, Windows 95 automatically recognizes that this is a word processing file and opens the word processing program. Windows 95 takes up about 20 MB of hard disk storage and uses 8 MB of RAM.

1.2.2.5.4 Windows NT:

Windows NT is another powerful operating system from Microsoft with multitasking and multiprocessing capabilities. It processes data in 32-bit chunks (unlike earlier versions of Windows, which process data in 16-bit chunks), resulting in increased speed and efficiency. Windows NT is ideal for large business applications that run in a networked environment; it provides mainframe like capabilities on a microcomputer. It can support multiple processors and has excellent I/O device support. Windows NT can

run on AT486 PCs. It requires 20 MB of RAM and occupies 40 to 45 MB of disk space. NFL selected Windows NT as its operating system because of its many features.

1.2.2.5.5 UNIX:

Although AT & Ts Bell Labs developed the UNIX operating system in 1969, it is only in the last decade or so that it has become popular. Today it is widely used in a number of important business applications. UNIX is a powerful, interactive and multiprogramming, multitasking (allows users to perform multiple tasks simultaneously), multiuser operating system that is highly portable (i.e. it works on a number and types of computers).

Wells Fargo Bank's wholesale services group, a company that provides cash management and related services to large U.S. businesses and financial institutions, chose UNIX as its operating system because of its portability and reliability. Applications dealing with the groups, core businesses total around \$20 billion in daily transactions run under UNIX. Another UNIX user is Western Publishing.

One of the disadvantages of UNIX is that there are many different versions of UNIX and this can sometimes get confusing. In addition, UNIX, compared to other operating systems, is cryptic and not very user-friendly. However, its advantages far outweigh its drawbacks and UNIX has become a mainstream operating system for many businesses.

1.2.2.5.6 OS/2:

IBM's OS/2 (Operating System/2) is 32-bit operating system that supports multitasking and can run programs written for OS/2 as well as for other operating systems, such as DOS and Microsoft Windows, thus reducing the need to learn several operating systems. Its 32-bit capability makes it faster than DOS. It is an ideal and sophisticated operating system for application, that requires networking and multimedia features, such as playing sound files or movies. OS/2 offers a number of small applications, called applets, such as time scheduling, appointment calendars, and card games. The OS/2 version of Windows has most of the features found in Windows 95, yet it requires only 4 megabytes of RAM.

In late 1994, IBM introduced its long-awaited new version of OS/2, dubbed OS/2 Warp. Since a dearth of brand-name applications that can run on OS/2 has always been OS/2's Achilles heel, IBM has effectively addressed the problem by ensuring, there are more than 2,500 applications that run on OS/2. Further, Warp can also run all the applications written for DOS and Windows thus greatly increasing the number of applications available to users who choose the OS/2 operating system. OS/2 Warp comes bundled with OS/2 applications, collectively known as Bonus Pak, which includes a word processor, spreadsheet, personal information manager, and easy access to the Internet and other on-line services.

In summary, operating systems are the most important type of system control software; they perform a number of basic functions. The following table provides a brief summary of the different types of operating systems. Clearly, no one operating system is superior to another, since a number of factors, such as number and criticality of applications, number of users, and network requirements, must be taken into account in selecting an operating system. Also, note that it is possible to use more than one operating system on a computer. A user may run some applications using, say, the

Windows operating system, and other applications under some other operating system, such as UNIX.

Table: Popular Operating Systems and their characteristics

Operating System	Characteristics
DOS	<ul style="list-style-type: none">▪ A single user system.▪ Commands are somewhat cryptic and not very user-friendly.
Windows	<ul style="list-style-type: none">▪ Uses icons to execute programs.▪ Uses point-and-click method to execute commands.▪ Applications appear in windows on the monitor.▪ A highly integrated environment, in which different applications have the same "look and feel".
Windows NT	<ul style="list-style-type: none">▪ Has multitasking and multiprocessing capabilities.▪ Processes data in 32-bit chunks.▪ Ideal for large business applications that run in a network environment.
UNIX	<ul style="list-style-type: none">▪ Powerful, interactive, multiprogramming, multitasking, multiuser, operating system.▪ Highly portable.
OS/2	<ul style="list-style-type: none">▪ This 32-bit operating system supports.▪ Multitasking, making it faster than DOS. <p>It can run programs written for OS/2 or for other operating systems such as DOS and Windows.</p>

1.2.2 System Support Software

System support software is a software that supports, or facilitates, the smooth and efficient operation of a computer. There are four major categories of systems support software: utility programs, language translators, database management systems, and performance statistics software.

Utility programs are among the most popular types of system support software. They perform tasks such as formatting disks, locating free space on a disk, retrieving lost or damaged files, sorting and merging data, converting files from one format to another, backing up important files, and providing online help.

As you may recall, all programs must be translated into machine language for the computer to understand and execute them. Language translators are support programs written specifically to convert a programming language (assembly language, high-level language, or very high-level language) into machine language. As you may recall, language translators that translate assembly language into machine language are called assemblers. Language translators that translate an entire program written in 3GL or 4GL into machine language before program execution are called compilers; language translators that translate a statement and execute it before translating the next statement are called interpreters.

Data is at the heart of an information system. Database management systems help us to create, manage, and maintain data. The operating system has certain basic data management features. Database management systems are unimportant system support software. Another example of system support software for organization-wide systems is

software that monitors the performance of a system by collecting data about various computer activities such as an idle CPU time: utilization of different I/O devices, amount of memory used by programs, and amount of time users are logged onto the system. By collecting data on system performance, managers can take action to ensure full utilization of their systems.

Most computer systems have support software, called utility programs, which perform routine tasks. These programs sort and copy data from one storage medium to another, output data from a storage medium to the printer and perform the tasks. The computer manufacturer as part of the operating system usually supplies utility programs. These utilities program may be called as applications program and are used by OS program.

Another common type of support software is a librarian. The primary function of the librarian is to maintain a catalogue of the locations and usage of all program and data files. Librarians often execute password controls.

Performance of monitors such as IBM's system management facilities (SMF) are a part of most system software. Performance monitors collect and record selected activities that occur within a computer system. For example, they collect data about CPU idle time, which operations are using the system and how long they use it and what hardware they employ, whether each job is successfully executed, and the amount of primary storage each job employs. This information can be used in charging departments within the firm for use of the computer facility. Most firms feel that charging users for computer services is an important part of the control over computer resources.

Monitors also collect information about which files are used in performing a job. This provides an excellent audit trail concerning data and file usage. It is possible to determine, for example, which files were used when a particular program was run. It will also identify, who was the user and when the file access was made plus the date and time of the access.

1.3.3 Systems Development Software

System development programs assist a programmer or user in developing and using an application program. Examples of these programs are language translators, linkage editors, and application generators.

A language translator is a computer program that converts a program written in a procedural language into machine language that can be directly executed by the computer. (Computers can execute only machine-language programs. Programs written in any other language must be translated into a machine-language load-module, which is suitable for loading directly into primary storage). Many different language translators exist; in fact, there is one for each programming language. They are categorized as compilers, interpreters or assemblers.

Quite often in writing a program, a programmer will call prewritten subroutines (or subprograms) which are stored on the system residence device, to perform a specific standard function. For example, if a program required the calculation of a square root, the programmer would not write a special program. He would simply call a square-root subroutine to be used in the program. The function of the linkage editor is to gather all of these called subroutines and place them into the application program. The output from the linkage editor is called a load module. The term module is often used synonymously

with program. A load module is a program that is suitable for loading directly into primary storage for execution.

Application generators are programming productivity tools that partially automate the programming process. For example, one type of application generator, the report generator, allows a programmer or user to simply describe the contents of a report rather than write the report in a procedural language such as COBOL. Some report generators (or query languages) even allow report requests to be made in conversational English.

1.4 Summary

In this unit, we have discussed about conceptual view of software and system software. We have also explained about kind of system software, which are more useful in application part of computer. Further, we have given more explanation about different type of operating system. We have identified various features of operating system during selection of proper operating systems.

1.5 Model Question:

Write a short note on the following.

1. What is system software? Can you define it with suitable example?

2. What do you understand by processing mode? Explain the different type of environment.

3. What is operating system? Write the various kinds of operating systems.

UNIT-2: SOFTWARE SOLUTION OVERVIEW AND APPLICATION

Structure:

- 2.0 Objectives**
- 2.1 Introduction**
- 2.2 Booting Loader**
- 2.3 Diagnostic Tests**
- 2.4 Operating System Executive**
- 2.5 Basic Input/Output system (BIOS)**
- 2.6 Utility Programs**
- 2.7 File Maintenance**
- 2.8 Application Software**
 - 2.8.1 General-Purpose Software**
 - 2.8.2 Dedicated Application Software**
- 2.9 Concept of Program and Instructions**
 - 2.9.1 Entering and Running Programs**
 - 2.9.2 Stored Program Concept**
- 2.10 Summary**
- 2.11 Model Questions**

2.0 OBJECTIVES

At end of this unit, candidate will be able to:

- Understand about the basic concept of software.
- Explain the various kind of general-purpose software.
- Know the concept various types of utilities.
- Analyse statistics of various file maintenance.
- Distinguish in between general-purpose system and application-dedicated software.
- Generate awareness about concept of program.
- Identify conceptual view of stored program.

2.1 INTRODUCTION

Software plays a very important role in computer field, because in absence of software, hardware has no meaning and without software we cannot perform any operation. So, the software is only part by which, we can apply same computer in different field. Computer is in all-rounder device and can be applied in different field.

2.2 Booting Loader

The function of the bootstrap loader program, which is usually contained in the ROM is to read the main portion of the operating system from secondary memory when the computer is first turned on.

2.3 Diagnostic Tests

The operating system contains a number of diagnostic examples; one diagnostic program might check the operable RAM. Some of these diagnostic programs are contained, turned on. Other diagnostic programs may second the operation of the system components.

2.4 Operating System Executive

The operating system executive is a program that controls the activities of the system, including executing programs and assigning tasks to hardware devices. In response to user or operator requests, the operating system executive schedules the execution of programs. The operating system executive is read into RAM when the computer first starts. Some sophisticated operating system executives (mostly on minicomputers and mainframe computers) allow several programs to be run simultaneously. This operating system feature is

called multitasking. Other operating system executives, called multi-user operating systems, allow several users to have simultaneous access to system devices and data.

2.5 Basic Input/Output system (BIOS)

Basic Input/Output system (BIOS) programs consist of the low-level routines used to control the most rudimentary functions of the hardware devices. Some functions of BIOS programs control are:

1. Reading the character from the keyboard.
2. Writing a character on the video display.
3. Writing character to the printer.
4. Determining whether the printer is busy or not.
5. Reading and writing a sector from the disk in a particular disk drive.

The BIOS programs are usually recorded in ROM. They are thus permanently available to programmers and can be used by programmers to create, write, read, and erase data. The details of BIOS program and its handling can be seen in the next unit.

2.6 Utility Programs

Every computer system needs utility programs to perform "housekeeping" functions. Some of the functions controlled by utility programs are:

1. Formatting a disk.
2. Displaying the contents of a disk.
3. Copying the contents of one disk onto another.
4. Determining the space remaining on a disk.
5. Making a backup copy of the contents of a hard disk.
6. Restoring the contents of a hard disk from a backup copy.

2.7 File Maintenance

The operating system provides service routines that can be used to maintain files on the various secondary memory devices attached to the system. These file maintenance programs allow user programs to create, read, and write files. These routines are not stand alone utility programs but are designed to be called from within user programs.

System development software helps system developers design and build better systems. An example is computer-aided software engineering (or CASE), a collection of programs that assist developers in developing an information system. Shells are also system development software; a shell is a set of tools and techniques that allows developers to build a prototype (or model) of a system. A shell contains some essential system features that have customized to meet the needs of the user. For example, shells have used to build expert systems, decision support systems, and executive information systems.

2.8 Application Software

An application program is a program written for, or by, a user to perform a particular job. General-purpose application software, such as an electronic spreadsheet, has a wide variety of applications. Specific-purpose application software, such as payroll and sales analysis is used only for the application for which it is designed.

Generally, computer users interact with application software. Application and system software act as interfaces between users and computer hardware. That software did not exist, very few people would have been using computer hardware. Due to involvement application and system software, computers become easier to use in different field. The Users Application Software Example: Payroll Program

System software controls the execution of the application software and provides other support functions, such as data storage. The MS-DOS, the computer's operating system, handles the basic platform. For example, when, you use an electronic spreadsheet on an IBM PC, the storage of the worksheet is file onto the disk.

Application software has been designed to perform people-related task such as payroll, inventory, and sales analysis. The United Nations uses application software to meet its information needs. There are two types of application software, general purpose (designed for general applications, such as payroll and so on) and dedicated software (designed for specific applications, such as the space shuttle). Different types of general-purpose application software groups considered below are listed in the following table.

- Word processing
- Desktop publishing
- Spreadsheets
- Computer graphics
- Accounting software
- Imaging

Type of Application	Software Package	Vendor	Operating System
Word Processors	Word	Microsoft	DOS, Windows
	WordPerfect	WordPerfect	DOS, OS/2, Windows
	Ami Pro	Lotus	DOS, Windows
Spreadsheets	Lotus 1-2-3	Lotus	DOS, OS/2
	Excel	Microsoft	Windows
Graphics	Power Point	Microsoft	OS/2, Windows
	Freelance Graphics	Lotus	OS/2, Windows
	Photo Shop	Adobe	Windows
Desktop Publishing	Publisher	Microsoft	DOS, Windows
	Print shop	BROAD	DOS, Windows
Accounting	Peachtree Accounting	Peachtree	DOS, Windows
	In the Black	Microfilms	Windows
	DacEasy Accounting	DacEasy	Windows
	MYOB Accounting	Teleware	Windows
	Simply Money	Kiplinger	Windows
Integrated Packages*	Microsoft Works	Microsoft	DOS, Windows
	Claris Works	Claris	Windows
*Packages that combine a number of capabilities, such as word processing, spreadsheets, and graphics.			

Table: Application software packages and supported operating systems.

2.8.1 General-Purpose Software:

General-purpose software has mass produced for a broad range of common business applications such as word processing graphics, payroll, and accounting. Some popular general-purpose software programs are to be discussed below.

Word processing is a computerized way to perform a task that has existed for centuries: text creation and management. It has been used by many organizations all over the world to save time and money and facilitate communication. Word processing programs allow for easy correction of errors

and text manipulation. Microsoft Word and WordPerfect are two very popular word processing packages.

Desktop publishing (DTP) has many of the features found in word processing packages and is used to produce office documents such as memos, price sheets, technical manuals, invoices, and new letters. Desktop publishing packages have more sophisticated graphical feature than do word processing packages. They have excellent text manipulation capabilities and color presentation graphics; these features can be used to produce high-quality documents. Desktop publishing provides WYSIWYG (What You See Is What You Get), which allows the user to see on the screen what the document will look like when it is printed.

Another type of general-purpose software is **spreadsheets**, which have been used for tasks that require number crunching such as financial analysis, budget preparation, and grade calculation. Spreadsheets can sort and analyze data, create charts, graphs, and figures, and perform "what-if" analysis to assess the impacts of changes to input variables.

The first electronic spreadsheet was **VisiCalc**, introduced in 1978 for Apple II computers. Since then, the spreadsheet has become an essential tool for professionals. In the early 1980s, Lotus introduced its versatile, easy-to-use Lotus 1-2-3 for the IBM PC. Today, there are other powerful spreadsheet packages on the market, such as Microsoft's Excel.

Although many word processing and spreadsheet packages are capable of producing **graphics**, sometimes users need sophisticated tools to draw complex figures and graphs. Computer graphics is the graphical display of computerized information; there are many sophisticated graphics packages. For example, Corel DRAW, a popular graphics program, offers many advantages over the more basic graphics packages included in word processing applications. Corel DRAW offers extensive text-handling and precision-drawing features; powerful paint and photo retouching applications featuring numerous image-enhancing filters, and even animation programs that let you create both simple and complex animations.

Accounting software is another type of general-purpose software, which helps companies automate their accounting functions. The main modules of an accounting system are General Ledger, Accounts Payable, Accounts Receivable, and Fixed Assets. These can be interfaced with other financial products, such as payroll, budgeting, and inventory management software to produce an integrated accounting package.

For example, Globex is an electronic futures trading network that uses accounting software to link brokers in New York, Chicago, and Paris with the Chicago Mercantile Exchange and the Chicago Board of Trade. Traders use this program to electronically buy and sell financial products and to receive timely and accurate financial information on various transactions. Accounting software

packages are flexible; they can be use for other applications. Budget Rent-A-Car uses an accounting software package to keep track of tosses resulting from damaged or wrecked vehicles.

Imaging Software uses scanners to scan paper-based documents, and then converts the data into digital images. Imaging software is increasing in popularity because of its many benefits. These include reduction of floor space when fewer paper documents are stored, decreases in copying costs, and fewer problems with lost or misplaced data. In fact, imaging software is becoming so popular that business analysts forecast that by 1996 imaging will be a \$12 billion to \$15 billion market in the U.S. alone; the worldwide figures are expected to be much higher.

2.8.2 Application-Dedicated Software:

The second type of application software is application-dedicated software, which includes specialized or customized applications designed for very specific purposes. For example, software designed to cater the delivery of goods and services to a customer's doorstep is highly specialized. Such a program cannot easily be modified and adapted for other applications because it is designed to perform a specific task.

2.9 Concept of Program and Instructions

A program is a sequence of instructions to be executed by the computer. Most computers are able to respond to a varied set of instructions, including numerical operations such as addition, subtraction, multiplication, and division; scientific functions usually found on a calculator; and text operations using letters, words, sentences, and paragraphs. The text operations that can be performed include arranging words as in an alphabetical order; scanning text for a particular word or phrase; displaying text on the screen or printer; or storing text in a long-term memory.

The computer can also make comparisons, such as: Is A greater than B? Is X equal to Y? Does the word alpha precede the word beta in alphabetical order? After making such a comparison, the computer can make a decision based on the result. The program can instruct the computer to do one thing if the comparison produces one answer or another thing. If the comparison produces the another answer, for example, an instruction may read

If $X > Y$ then print "X is larger" on the screen. If X is not larger than Y, the program will proceed to execute the next instruction.

Sometime a program will direct the computer to execute a group of instructions repeatedly. This is called a loop. Computer programs may instruct the computer to save data in the computer memory or on the disk. Some instructions or commands are given to the computer to retrieve a piece of data from memory. In some programs, the computer is instructed to send data to another location via telephone lines.

2.9.1 Entering and Running Programs

To be carried out a program must be placed in the primary memory of the computer. When not being used a program is stored in a long-term memory device such as disk or tape. When the program is to be used the computer is given appropriate command to read the program from long-term memory like floppy diskette into primary memory. On some computers, the user must give a separate command to execute the program when the computer carries out the instructions of a program. We say that the computer runs or executes the program. Executing the program involves both reading the program into primary memory and then carrying out the instructions of the program. Unless the program states otherwise, instructions in the program are executed in order.

2.9.2 Stored Program Concept

Since a computer can perform arithmetic operations, it may seem to be just a calculator. But, there is at least one important difference with a calculator, it is necessary to enter instructions one at a time. With a computer, the complete sets of instructions are stored in the primary memory of the computer. The idea of storing the program into primary memory of the machine is called the stored program concept and is fundamental to all modern computers.

It is difficult to overemphasize the importance of the stored program concept. With a calculator, if you wish to perform a set of instructions many times, it is necessary to enter them repeatedly. With the computer, since the instructions are stored in memory, you may re-execute them without reentering them. Once it is determined that a program is working correctly, the computer may be set to work performing the instructions by itself. The computer works with incredible speed and with good accuracy. It is easier to make mistakes with a calculator, especially if you are performing calculations involving thousand of operations.

As we know, the computer is a machine, which processes data into a useful form, information. But it requires the sequenced instructions to do so. A computer cannot provide its own input. It cannot decide how it can be programmed and cannot implement any decision on its own; and it cannot think. It needs to be programmed. Therefore, the most import component of a computer system is its live ware.

System programs are developed and sold by the computer companies and specialized software firms. System programmers write system software. Most large firms have their own systems programming staffs that are capable of modifying an operating system to meet the unique requirements of the firm.

2.10 Summary

In this unit, we have discussed about conceptual view of software. We have also explained about kind of application software, which are more useful in application part of a computer. Further, we have given more explanation about different steps of utilities. We have also explained on the various type of general-purpose software used in computer field.

2.11 Model Question:

Write a short note on the following.

1. What is software? Can you define types of application software?

2. What do you understand by file maintenance? Explain the different type software used in computer field.

3. What is meant of general-purpose software? Differentiate it with application dedicated software.

UNIT- 3: BASIC INPUT OUTPUT SYSTEM HANDLING

Structure:

- 3.0 Objectives
- 3.1 Introduction
- 3.2 BIOS Basics
- 3.3 PC-BIOS Communication
- 3.4 Other BIOS Functions
 - 3.4.1 POST (Power On Self Test)
 - 3.4.2 Bootstrap loader
- 3.5 BIOS Version
- 3.6 BIOS Compatibility
- 3.7 System Identification
- 3.8 BIOS Limitation
- 3.9 BIOS Extension
- 3.10 Upgrading BIOS
 - 3.10.1 System Upgradation
 - 3.10.2 BIOS Modification
 - 3.10.3 Copying BIOS
- 3.11 Flash BIOS
- 3.12 IML Scheme of IBM
- 3.13 Summary
- 3.14 Model Questions

3.0 OBJECTIVES

At end of this unit, student will be able to:

- Understand about the basic concept of BIOS.
- Explain the various role of Basic Input Output System.
- Know the concept of POST and Bootstrap Loader.
- Analyze statistics of BIOS limitation.
- Discuss the upgrading and modifying BIOS in the computer system.
- Identify conceptual view of FLASH BIOS.

3.1 INTRODUCTION

BIOS or the Basic Input Output System is a set of programs stored inside a Programmable Read Only Memory (PROM) and put into the motherboard. The main job of the program is to load into memory and used facilities to provide the user with a set of standard routines to take care of input/output from different input, output and storage devices connected to the computer.

This BIOS ROM is always available in the computer, so any user program can access these routines for their input/ output requirements.

The IBM for their PC range of computer made the first PC BIOS. The BIOS program is being the copyright property of the IBM Corporation. PC compatible makers use the BIOS made by other companies such as AMI, Award and Phoenix etc.

The BIOS made by these compatible makers intimate the original IBM BIOS without using the same program code.

3.2 BIOS Basics

BIOS as the name suggests is used for most basic input output process such as reading the keystroke, displaying something on the screen, reading / writing of the serial and parallel port, reading / writing to and from the floppy and hard disk etc.

The BIOS is a set of programs or software stored inside a PROM, which is hardware, so BIOS is not complete software or a complete hardware. A new name "firmware" is used to describe the BIOS. The term firmware is used to specify any ROM or PROM based on software or program.

The external device is connected to some I/O port know to the BIOS. When the computer or the user program asks the BIOS to read the keyboard it is the function of the BIOS to read the right I/O port in a proper manner, so as the data from the keyboard can be made available to the computer or the user program.

If you have two different computers one having keyboard connection at one I/O port 60 and the other computer having the keyboard connection at I/O port 62, you do not have to purchase a different program for each of them.

The computer or your program can just give the read keyboard command to the BIOS and it will be the job of the BIOS to read proper I/O port 60 for the first computer and port 62 for the second computer.

3.3 PC-BIOS Communication

The computer or user program and the BIOS communicate with each other using something called interrupts. Whenever the computer has to request for some information to the BIOS, computer generates an interrupt, and the BIOS will carry out a specific work such as reading the keyboard or displaying something on screen etc according to the interrupt number. These interrupts are software interrupts and are different from the hardware interrupt generate by some devices to communicate to the processor. Just like a hardware interrupt, when the processor receives a software interrupt, the processor stops whatever it was doing and takes care of the interrupt.

When an interrupt is received, the BIOS will search the interrupt vector table for the address of the program that should provide the required service to the interrupting program.

Once the address is found, BIOS starts execution of the interrupt service routine (a program that provides the requested service to the interrupting program). After finishing the interrupt service routine the BIOS will return to the old process that has initiated or called the interrupt routine.

The interrupt vector table is located at the lowest RAM memory address that is at address 0000 at segment 0000 and contains 4byte for each interrupt number.

This table is loaded from the BIOS ROM to the RAM at the time of system startup and loading the address from the ROM to the RAM. It is done so that, if the user wants to provide some other interrupt service routine than the routine provided by the BIOS, he can do so by changing the routine address in the RAM vector table.

3.4 Other BIOS Functions

Other than the input/ output function explained so far the BIOS contains two more modules

- Power on self test and
- Bootstrap loader

3.4.1 POST (Power On Self Test)

POST or Power On Self Test is another program stored inside the BIOS ROM, the basic job of this program as the name suggest is to check the processor, memory, support circuits and other devices connected to the computer when the computer is first powered on or switched on.

When the computer is switched on the POST routine inside the BIOS ROM takes over and checks all the basic devices connected to the computer, if this routine finds any error or fault during the POST routine the error depending on its severity is informed to the user as error beep code or as some error message displayed on the screen.

Some errors called non-fatal errors allow the user to continue the process, whereas the fatal errors will not allow the POST to continue to the next step until the problem or the fault is rectified.

Once the POST routine is executed successfully, BIOS goes to the next process that is the boot process.

3.4.2 Bootstrap loader

The next part of the BIOS, after the POST routine contains a program called bootstrap loader. It is a function of this part of the BIOS to load the operating system from the floppy disk or the hard disk into the computers main memory (RAM).

Unless the boot process is successfully completed i.e. the Operating System (OS) is copied to the main memory the user cannot use the computer. A successful transfer of the OS into the RAM or the boot process is indicated by displaying the DOS prompt such as A :> or C :> on the computer monitor. Once this prompt is displayed on the screen, the user can start using the computer.

If you have Window installed into your system then instead of the DOS prompt, you will directly go to the Window screen after a successful boot operation.

3.5 BIOS Version

Different version of the BIOS provides different capabilities. Sometime if you want to install a new device to your computer, your BIOS should have the support for that particular device.

For example, many old BIOS do not have support for 2.88MB 3.5 inch floppy drive. Sometimes some old versions of the BIOS may contain some error or bug such as the AMI BIOS before April 1990 did not work properly with the AT interface hard disk drives.

If you want to check the revision date of your BIOS, you can use the following BASIC program or the DOS DEBUG script as explained below. Load the GWBASIC and type the next BASIC program. After typing the program without any mistake, type RUN command and press Enter key.

This will display the revision date of your BIOS with some additional characters.

```
10 DEF SEG=&HF000
20 FOR I = &HFFF5 TO &HFFFF
30 PRINT CHR$(PEEK (I));
40 NEXT I
50 END
RUN
```

Or

Load the DOS DEBUG program and type the below given DEBUG script to check the revision date of your BIOS.

```
C:\DOS62>DEBUG
```

```
-D F000:FFF0
```

```
F000: FFF0 EA 5B E0 00 F0 30 36 2F-30 36 2F 39 32 00 FC 00
```

```
.... [.....10/10/2007...
```

```
-Q
```

```
C:\DOS62>
```

3.6 BIOS Compatibility

The standard BIOS or the IBM BIOS is available with the IBM computers only. Most of the compatible manufacturer use BIOS made by many different IBM BIOS compatible manufacturers.

As the copyright law protects the original IBM BIOS, the compatible makers try to give the same routines by writing their own code.

To do this without breaking the copyright laws, the programmers study the input and output of different routines provided by the IBM BIOS and duplicates the same without looking at or copying the IBM BIOS routine. This gives a different program than the BIOS program made by the IBM, but the input and output from this routine will be same as the original IBM BIOS.

Sometimes some program that directly use the IBM BIOS i.e. instead of using the interrupt, directly call the BIOS routine may not work on these compatible BIOS as the program of the original IBM BIOS and the compatible BIOS are very much different.

3.7 System Identification

The BIOS also contains a system identification code at the memory location 0FFFFE, the user program to find out the type of the system being used to run the program can use this code. This code is IBM's system identification code, but most of the compatible makers use this byte value to identify their system as well.

You can use the DOS DEBUG program or the GWBASIC program given below to find out the type of your system. Load the GWBASIC and type the next BASIC program. After typing, type RUN command and press Enter key. This program will display the type of your system on the screen.

```
100 DEF SEG = &HF000: C% = PEEK (&HFFFE)
110 PRINT "Type of the system: "
120 IF C% = 248 THEN PRINT "PS/2 Model 80"
130 IF C% = 249 THEN PRINT "PC Convertible"
140 IF C% = 250 THEN PRINT "PS/2 Model 30"
150 IF C% = 251 THEN PRINT "PC/XT or Compatible"
160 IF C% = 252 THEN PRINT "PC/AT or Compatible"
170 IF C% = 253 THEN PRINT "IBM PC Junior"
```

```

180 IF C% = 254 THEN PRINT "PC/XT or Compatible"
190 IF C% = 255 THEN PRINT "IBM PC or XT"
200 IF C% < 248 THEN PRINT "Unknown System ID:" HEX$(C %); " Hexa"
210 END
RUN

```

Load the DOS DEBUG program and type the below given DEBUG script to check the System Identification Code of your system.

```

C:\DOS62>DEBUG
-D F000: FFFE
F000: FFFE FC 00...
-Q
C: \DOS62>

```

3.8 BIOS Limitation

Even though, the BIOS is made to take care of a large number of different devices and activities. It has some limitations. One major limitation of the BIOS is as it is not made for some specific machine, using the BIOS routine for the I/O purpose restrict the speed and the facilities provided by some devices.

For example, when you want to output some character on the screen, you need to first move that character and their attribute and screen location etc. into some processor register. Next, you need to call a BIOS routine using proper interrupt number to output this information onto the screen.

The BIOS routine will again read this information from the registers and move them to the proper display memory location to make the character visible on the screen.

This whole process makes the screen display very slow when it is done through the BIOS. This has made most of the software developers to directly write the screen memory bypassing the BIOS for a fast screen display.

This direct hardware control by bypassing the BIOS routine is a common method to gain additional speed in many operations by the software developers. Sometimes this is also done to use some hardware for a purpose other than allowed by the BIOS, for example to make a copy protected floppy disk one can format a floppy in some format that can not be read by the standard BIOS routines.

The direct hardware control has one major shortcoming, as the memory or the port used by different devices need not be the same on the entire computer, the program that directly writes to the memory or the I/O port, bypassing the BIOS, may not work on some compatible computers.

Whatever be the BIOS limitation, it provides the programmers with a standard set of software routines to communicate with the devices connected to the computer.

Without the BIOS, each programmer will have to write program to read and write each device connected to the computer.

3.9 BIOS Extension

As the BIOS does not provide software drivers for all the devices that one may connect to the computer. The extension card used to connect different devices to the computer contains additional ROM modules.

The PROMs contain device specific routines that user programs can call to interact with the device connected to the extension card.

When the system is switched on, after completing the POST process and after copying the BIOS interrupt vectors to the RAM, the BIOS program checks all the memory location 0C8000 to 0F4000 in 2K blocks for any additional ROMs. The system designers for the ROM extension set this memory aside.

If any valid ROM is found in that location, the program in the ROM will be added to the basic BIOS routine and the interrupt vector is changed accordingly.

To check for a valid ROM, the BIOS searches memory location 0C8000 to 0F4000 for a special two byte code 055(Hex) followed by 0AA(Hex). Once these two codes are found the third byte is checked. Third byte contains the length other BIOS extension program in number of 512-byte long blocks.

Once the length of the program is found, BIOS does a Cyclic Redundancy Check (CRC) on the specified number of blocks. If no error is found in the CRC check, it indicates a valid BIOS extension module at that particular memory location.

When a valid adapter BIOS is found BIOS starts executing it from the fourth byte onwards using a FARCALL instruction. Mostly this code initializes the device for which it is written and adds or modifies the interrupt vector for further communication by the other user programs. After this, the adapter BIOS gives the control back to the main BIOS.

The main BIOS will continue to search for the valid adapter BIOS until memory location 0F4000 is reached, after that it will start the boot process.

This check allows multiple adapter BIOS ROMs to be used in the PC system. Only limitation is that no two devices can share the same memory area, most of the extension cards with ROM have some type of setup facility, DIPswitch or jumper to set the memory address to be used by them. Avoiding this type of conflict is covered in details in motherboard upgrade unit under the system configuration section.

3.10 Upgrading BIOS

You may want to upgrade your existing BIOS, if it does not support the device, which you want to use in your system.

Some old BIOS do not have support for 1.44MB or 2.88MB 3.5 inch floppy disk drives, your existing BIOS may not support the hard disk or the display you want to use in your system. In these conditions, you have two options

- Either you can buy a new BIOS upgrade or

- You can modify your existing BIOS.

Buying a new BIOS and installing it into your system is an easy approach compared to the modification of the existing BIOS. You will need some additional hardware devices and a thorough knowledge of hardware and software to modify your existing BIOS.

3.10.1 System Upgradation

To upgrade your existing BIOS, you will require new BIOS, which should contain the facility, required by you and should be compatible to your existing hardware. It should not happen that after installing the new BIOS, some of the old devices in your machine stop working.

When buying a new BIOS chip you need to inform the sales person complete detail about the system for which you require the BIOS.

Some of the information you need to give are:

- Make and Model of the system and the motherboard.
- CPU type and speed.
- Make or part number of the motherboard chipset such as Chips technologies and Optic etc.
- Make and version of your existing BIOS ROM.
- Additional devices that you want to use in your system, such as VGA display 2.88 MB floppy disk drive etc.
- Make and version of the existing Keyboard Controller chip.

Once you have the required BIOS, you can upgrade it as you would upgrade a normal memory chip. Switch off the system, open the cover, remove the old BIOS, put the new BIOS into the socket, switch on the system, run the setup option and check if everything is working properly, before closing the cover.

The process for removing and putting the BIOS chips is same as the RAM chip upgrade, which is explained in the memory upgrade unit.

Keyboard Controller: Keyboard controller chip is a microprocessor with built-in ROM; this ROM is programmed by the BIOS manufacturer to read keyboard, control the reset and A20 line for protected mode operation etc.

The Intel 8042 microcontroller, which is a microprocessor with in-built RAM & ROM memory, and I/O ports, is most commonly used as keyboard controller chip.

This is a 40-pin DIP chip somewhere near the keyboard connector on the motherboard. Most of the time, this chip is covered with a copyright notice label, which identifies the manufacturer and version of the BIOS code stored into it.

As this chip is used for keyboard reading and protected mode operation, if the BIOS is not compatible with the keyboard controller you may face problems with the keyboard or protected mode operation.

This could result in not being able to load the (HIMEM.SYS) OS extended memory manager, which requires support for A20 address line.

This could also give problem with the Window or OS/2 software, which runs in protected mode.

If you face any of these problems with your new BIOS, then you may need to change the keyboard controller as well. On most of the motherboards, the keyboard controller is installed into socket, which makes it easy to upgrade.

3.10.2 BIOS Modification

Other than the BIOS upgrade you can also do the BIOS modification, if you have proper hardware and software knowledge and EPROM programming and erasure device.

Using these devices, you can make any change to the existing BIOS code. You can change the system start up message being displayed or the error codes displayed, or you can make change to support some specific device etc.

Whatever change you make, make sure that the change you make should not make the BIOS incompatible to the existing software; otherwise, you may not be able to use the standard PC software.

3.10.3 Copying BIOS

Before making any change to the existing BIOS, it is always safe to make a backup copy of your old BIOS code, as many a time you may not find the exact copy of your old BIOS, if something goes wrong during the change.

The EPROM programmers can read from the EPROM and write to the disk file, this facility can be used to make a disk backup of your BIOS code.

You can also copy the disk backup file to an empty EPROM using the EPROM programmer, as most of the EPROM programmers can read from the disk file and write to the EPROM.

3.11 Flash BIOS

As a ROM, based BIOS cannot be changed without going through the long process explained above and requires additional devices to program and erase its content, a new type of BIOS chips are gaining popularity with the new systems.

This new BIOS is called Flash BIOS. This is a particular type of EEPROM (Electrically Erasable Programmable Read Only Memory) chip that can be erased and programmed in the system itself; it does not require any Erasure or programming device.

Using a Flash BIOS in a system allows the manufacturer to upgrade or modify the existing BIOS code by sending an upgrade on a disk to the users. The user can then download the content of the floppy to the Flash BIOS without removing the BIOS from the system.

Normally all the Flash BIOS based system have some type of hardware write protection to the flash BIOS, otherwise a virus program may change or destroy the BIOS code and make the system crash. This write protection is implemented using some type of switch or jumper on the motherboard.

3.12 Initial Microcode Load Scheme of IBM

IBM uses a system similar to the Flash BIOS called Initial Microcode Load (IML) in some of their new PS/2 systems. In this system, they keep a very small program in the ROM on the motherboard, which will load the actual BIOS system from some hidden area of the hard disk into the computers RAM memory at the system start-up time.

In this method the actual BIOS is stored in a hidden protected 3MB partition on the hard disk, and this allows the IBM to send any BIOS upgrade on a reference disk to the users. The user can simply boot this new reference disk and choose the option to copy its content to the hidden partition; this will change the existing BIOS code with the new code provided in the reference disk.

This hidden partition does not appear when running the FDISK or the FORMAT command, which keeps it safe.

A small ROM program on the motherboard loads this IML BIOS to the system RAM at every time the system is switched on. This gives the impression as if the complete BIOS is inside the ROM. The IML also contains a setup program, which allows the user to change the system setup just as it is done on normal BIOS ROM based systems. One drawback of this system is if the hard disk fails, one can not boot the system using the standard boot diskette; it will require special reference floppy disk which IBM provides with the system.

3.13 Summary

In this way, we have seen that BIOS is more important equipment, which used to control input / output of devices of the computer. It is generally in ROM of the computer at the time of manufacture permanently. We have also discussed booting process of computer, which makes computer workable. We have also explained the upgradation of new BIOS in the system.

3.14 Model Questions

Write a short note of the following.

1. What do you understand about BIOS? Explain its role in computer.

2. What is role of Bootstrap loader in the computer? Explain it.

3. How you upgrade anew BIOS? Describe its steps.

4. Describe the limitation of BIOS.

UNIT- 4 : INSTALLING AND CONFIGURING THE DEVICE DRIVER

Structure:

- 4.0 Objectives**
- 4.1 Introduction**
- 4.2 Installing Device Drivers**
 - 4.2.1 Plug and Play**
 - 4.2.2 Installing New Hardware**
 - 4.2.3 Responding to a prompt for a Driver**
 - 4.2.4 Signed Drivers**
 - 4.2.5 When Drivers are unavailable**
 - 4.2.6 PnP giving a little help**
 - 4.2.7 Installing non-PnP Hardware**
- 4.3 Updating a Device Driver**
- 4.4 Rolling Back a Driver Update**
- 4.5 System Resources**
- 4.6 Working with Device Manager**
 - 4.6.1 Troubleshooting of Device Problems**
 - 4.6.2 Manual Handling the Device Resource**
 - 4.6.3 Removing a Device Driver**
- 4.7 Summary**
- 4.8 Model Questions**

4.0 OBJECTIVES

At end of this unit student will be able to:

- Understand about the basic concept of device drivers.
- Distinguish between PnP and non-PnP device drivers.
- Explain the various functions of device drivers.
- Know the concept of various types of device drivers.
- Analyze statistical feature of various device drivers in installation.
- Generate awareness about general trouble shootings of device drivers.
- Identify conceptual view of system development software.

4.1 INTRODUCTION

The hardware and software of a computer system are mutually dependent on each other. For a new piece of hardware to work, your software (particularly in your operating system) must be aware of the new device and must be configured appropriately. This means that whenever you are working on your system's hardware, you are also interfacing with the operating system, which introduces an extra level of complexity to the operation.

Because of this interrelationship between hardware and software, you need to be aware of and know how to use various configuration and system management utilities that are built into your operating system. If you know what you are doing, you can use these built-in Windows utilities to manage the installation and operation of your hardware and to track down any hardware-related problems that may develop.

4.2 Installing Device Drivers

Any piece of computer hardware is just a useless chunk of circuitry unless it can communicate with your PC's operating system. That communication happens through a device driver, which is a piece of software designed to serve as a translator interface between the operating system and the hardware. So, for example, if a certain piece of hardware is "Windows XP compatible", it means there is a Windows XP driver available for it.

Any hardware can work with Windows, given an appropriate driver and a way of physically connecting them and assigning system resources. You could even connect your refrigerator to a PC and talk to it, if the connectors and drivers were present. (You surprise, but that day is coming sooner than you think).

So what do we mean by the right device driver? There are two things in this connection:

- The driver is written for the exact make/model of device you have.
- The driver is written for the exactly synchronized version of Windows you have.

Sometimes you can feel your way through with a driver that does not meet both of these criteria but not always. So, when we talk about installing a piece of hardware in Windows, what we are really talking about is installing the correct driver for it. Plug and Play (PnP)

makes the process simple in many cases; in the rare cases where PnP does not work, you can fall back to manual techniques.

4.2.1 Plug and Play

PnP is a hardware identification system that enables the Basic Input Output System (BIOS) and the operating system to immediately see new hardware and assign resources to it. And, if a driver is available for the device, PnP will try to install it in the operating system.

PnP requires three things to work:

- The BIOS on the motherboard must support PnP.
- Most motherboards made of since 1995 or before.
- The device itself must support PnP.
- The operating system must support PnP. Windows 98 or higher version support. Windows 95 and NT4 do not support.

For a device to work in Windows, it has to have a compatible driver, of course, and this is a separate issue from whether it is PnP-capable. However, Windows (especially the recent versions) has built-in drivers (or drivers readily available on its Setup CD) for thousands of devices, so it may seem as if PnP is also providing a driver. It is not. The PnP process looks for the driver and automates its installation if found, but it does not actually provide the driver.

Therefore, during the PnP detection process for new hardware, you might be prompt to insert a CD containing the needed driver. If Windows has a suitable driver built-in, it installs it; otherwise, it prompts you.

4.2.2 Installing New Hardware

When PnP works flawlessly, you start up the PC, and Windows detects the device, it installs a driver for it and the device works. When it does not well, see the following sections.

4.2.3 Responding to a prompt for a Driver

When a prompt appears for a driver, it means Windows does not have one for the device. You have three choices:

- You can insert a disk or CD containing the needed driver. If it cannot find the file, a message will appear, and you might need to browse the disk or CD to locate the driver. For example, the CD that came with a device might have separate folders for each operating system version, and you might need to navigate to the correct folder.
- If you cannot insert a disk or CD and click OK, a message will appear telling you that it still cannot find the driver and allowing you to enter a path or browse for one. Browse to a location on your hard disk or other disk or CD that contains the needed driver. This works well when you have downloaded a driver to use.

- You can click Cancel to abort the PnP installation of the device and then run the setup software that came with the device. This is the best choice for devices that came with special software you need or want to use, such as a multifunction printer that has its own scan/ fax/copy/ print utility.

4.2.4 Signed Drivers

A few years ago, one of the most common ways, Windows could be screwed up was with third-party hardware to come with its own drivers that had not been tested well with all Windows versions. Windows 2000 and higher attempt to prevent such problems by preferring signed drivers. A signed driver is one that has a digital signature that certifies it has been tested with certain versions of Windows and proved to work, and that it has not been altered since the signature was granted. By installing only signed drivers, you can virtually eliminate driver based system problems.

When you attempt to install a certain driver in Windows 2000 and higher, Windows checks to see if it is signed, the installation proceeds. If it is not, you see a warning box. You have the option of allowing the unsigned driver to be installed or aborted. (That is, unless your network administrator has clamped down on your security settings and made it impossible to install unsigned drivers).

If you are installing an unsigned driver for a non-critical device such as a printer, scanner, or modem, has no big deal. Prefer the signed driver, but if you do not have one, try the unsigned driver. It will probably work. However, if you are installing an unsigned driver for some critical component such as the video card, beware. Back up your Registry first (using System Restore in Windows Me or XP or the Registry Editor in other versions), and do not do the install at all if there is any way you can possibly get your hands on a signed driver instead.

To control Windows' behaviour in the face of an unsigned driver, do the following (these steps are for Windows XP, but for Windows 2000 steps are similar).

1. From the Control Panel, double-click System.
2. Click the Hardware tab.
3. Click the Driver Signing button.
4. Choose a warning level, as in figure below
5. Click OK.

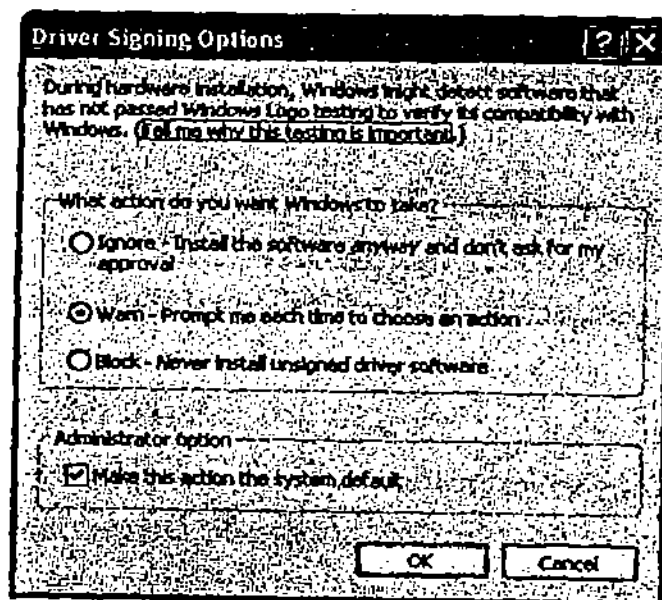


Figure: Changing the warning level for unsigned drivers

4.2.5 When Drivers are unavailable

The ideal driver is written for exactly your model of hardware and for exactly your Windows version. And ideally, you will have the latest one on a Setup CD that came with the device. The ideal does not always happen. Therefore, the following are some ways around a no-driver situation, in order of preference: Go to the device manufacturer's Web site and download the correct driver, making sure you get the latest one for your operating system and your hardware model. Unpack and install it as recommended by the manufacturer. (It might be a Zip file, but it will probably be a self extracting EXE or an EXE Setup program).

Check other Web sites that might have the driver you need, such as Win Drivers (WWW.windrivers.com). Using the Internet is devoted to finding the hardware support on it.

Back up your Registry first (using System Restore in Windows Me or XP or the Registry Editor in other versions) and do not install at all if there is any way you can possibly get your hands on a signed driver instead.

To control Windows' behavior in the face of an unsigned driver, do the following (these steps are for Windows XP, but steps are also similar in Windows 2000):

1. From the Control Panel, double-click System.
2. Click the Hardware tab.
3. Click the Driver Signing button.
4. Choose a warning level as per figure below.
5. Click OK.

Attention: In versions prior to Windows XP, sometimes the New Hardware Wizard will find the wrong things, such as a stray port that you have not been using--perhaps built-in

controller on the motherboard that is not connected anything or perhaps the PS/2 mouse port when you have been using a Universal Serial Bus (USB) mouse. So, before you allow it to install the drivers and assign resources for the devices it finds, you should review the list. If it found the wrong thing, Click Cancel.

If the wizard does not find the device, your next stop should be Device Manager, where you will see whether any trace of the device appears and figure out why Windows does not like it. We will explain that process later.

4.2.6 PnP giving a little help:

Sometimes Windows can detect that a new device is present, but it has no clue as to what type of device it is or what driver might be appropriate. One indicator of this would be the device showing up in Device Manager (covered shortly) as Unknown Device. When that happens, all you need to do is run the Setup program that came with the device. This tells Windows the device's make and model and installs a driver.

4.2.7 Installing non-PnP Hardware: The main thing about non-PnP is its hardware. The difference is that Windows cannot assign resources to it automatically. For example, if the device needs an Interrupt Request (IRQ), you must set the jumpers on the device to a particular IRQ. Windows has no say in the matter; it merely reads the IRQ setting on the device and then tries to shuffle around all the automatic assignments for other devices to accommodate it. If, it is possible or not, if not, the one or more devices may not work until you make some manual resource adjustments in Device Manager.

The best way to integrate a non-PnP device into a system is to let it have first pick up the resources. Install it first, and then boot into Windows and run the Add Hardware Wizard to cause Windows to detect it and read its resource claims. After the device is working, then install the PnP devices (which ostensibly can use whatever resources are left).

If you are installing a non-PnP device into a system that already has a lots of PnP devices, you might save some time by trying the non-PnP device without removing anything. It might work. If not, you can take a fallback position of removing PnP devices that are currently using the resources it wants and then reinstalling them later.

4.3 Updating a Device Driver

Hardware manufacturers periodically release new versions of their drivers for various operating systems, both to patch bugs and to introduce new features. You can check the manufacturer's Web site periodically and download whatever is available.

Attention: Some hardware maker's will let you register to receive e-mail notification whenever a new driver is available for a certain hardware model.

Once you get that new driver, it will be in one of these formats:

- Executable installation files with an EXE extension
- Executable self-extracting Zip file with an EXE extension
- Zip file (requires unzipping) containing the files needed to run a Setup utility
- Zip file containing bare files (such as DRV and INF) requiring manual installation

If you get an executable file, run it. If a Setup program runs greater, use it. If a WinZip self-extractor appears, extract the files to some temporary location on your hard disk and then check out what you got. If there is a Setup.exe file in there somewhere, run it.

The same goes for getting a ZIP file, except you will need an unzipping utility such as WinZip to extract the files (unless, of course, you have Windows XP or higher, with built-in ZIP support.) Place the files in some temporary location on your hard disk, and again, check out what you have. Run the SETUP.EXE if there is one. We have saved the hardest possibility for last, no Setup utility, just a bunch of files with obvious way for installing them. If that is what you have ended up with, you will need to do the following:

- Open Device Manager. For Windows 95/98/Me, open the Control Panel, double-click System, and click the Device Manager tab. Similarly, for Windows 2000/XP/2003, open the Control Panel. Double-click System; click the Hardware tab, and click the Device Manager Button.
- In the list of device types, click the plus sign next to the type of the desired device. That device's name appears on the expanded list.
- Double-click the device name to see its Properties box.
- Click the Driver tab.
- Examine the details of the current driver, noting the date and whether it is signed. Proceed only if the driver you are replacing it with is a better one (in other words, a more recent one or signed when the current one is not).
- Click the Update Driver button. The Hardware Update Wizard runs. Follow its prompts to install the new driver. (It will ask you for the location at some point; browse for it.)

4.4 Rolling Back a Driver Update

After updating a driver, you can revert to the earlier version with the Roll Back Driver feature:

1. Open Device Manager and display the Properties box for the device.
2. On the Driver tab and click the Roll Back Driver button. Then follow the prompts.

4.5 System Resources

Each hardware device that interacts with the central processing unit (CPU) needs resources assigned to it. Different devices have different resource requirements, but they all need at least one of these types:

IRQs are signaling lines between the CPU and the device. The devices use the IRQ line to signal to the CPU that it wants to have a conversation. IRQs are numbered 0 through 15. Some devices—for example, Industry Standard Architecture (ISA) boards—require their own reserved IRQ; others—for example, Peripheral Component Interconnect (PCI) boards—can share.

Memory addresses These are areas of Random Access Memory (RAM) set aside for the device to use. Some devices need a range of memory reassigned to them; others can take from the general pool as needed.

I/O addresses Input/ Output (I/O) addresses are also areas of RAM, but these are reserved for the transfer of data to and from the device.

DMA channels Direct Memory Access (DMA) channels are less commonly needed on modern devices. Keyboards and sound cards used to need them, but not so much anymore. They are pathways that allow the device to read and write directly from memory, bypassing the CPU. They were helpful back in the days when the CPU was not very fast, but nowadays they are not that important.

Many of the most common device types have certain resources pre-allotted to them by default. For example, the legacy serial port COM1 has IRQ4 and I/O address 3F8-3FF. The PnP portion of the BIOS does, these pre-assignments at startup, and then when Windows starts up, it accepts them. If you take the A+ Certification exams, you will want to memorize the default assignments, but it is not necessary otherwise.

When two or more devices want the same resource, a conflict occurs. This is a real possibility if you have multiple non-PnP devices, but most of the PnP devices are flexible enough to accept wherever resources happen to be free. If a conflict among defaults does occur, Windows works out compromises among all PnP devices so that every device is satisfied. This all happens automatically and invisibly for the most part. However, situations may arise where Windows is unable to handle resource assignments completely independently and that is where you come in. You can use device Manager to straighten it all out.

4.6 Working with Device Manager

Device Manager is a hardware that Windows recognizes as installed on the system, broken down by category. You can select any device and view its status and properties from there, as well as manually change its resource assignments in some cases.

To access Device Manager, use these following steps:

- For Windows 95/98/Me, in the Control Panel, double-click System, and then click the device manager Tab. Device Manager appears as a tab in the dialog box.
- For Windows 2000/XP /2003 Server, in the Control Panel, double-click System, and then click the Hardware tab. On the Hardware tab, click the Device Manager Button. Device Manager appears as a separate window.

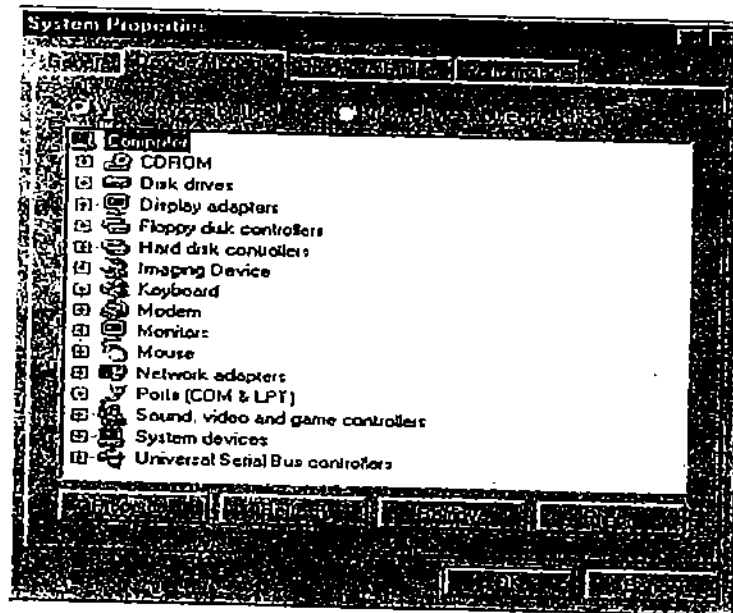


Figure: Viewing the Device Manager during installation.

Attention: A shortcut for displaying the System Properties is to right-click My Computer and chooses Properties.

Because every computer has a little different, chances may your list will not be look exactly like this one, but it will have most of the same device types. The information in the list is organized like an outline, listing categories of devices. The devices themselves appear within each category. To see the devices, click the plus sign beside the category.

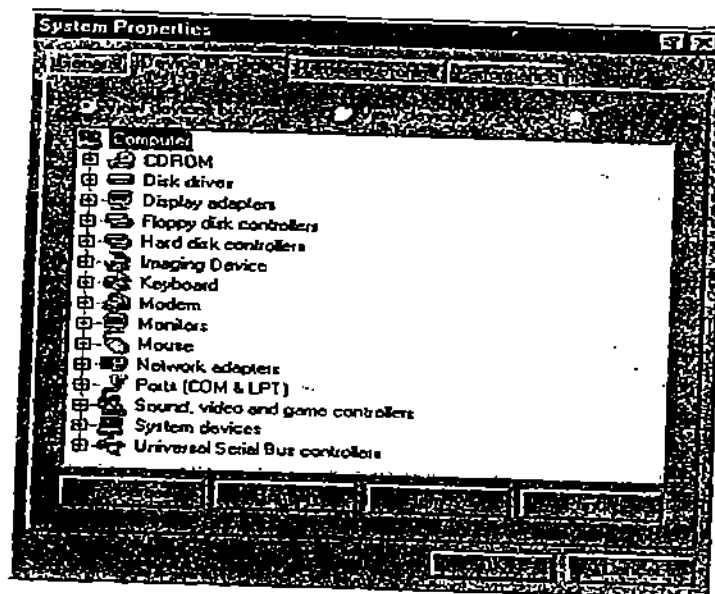


Figure: Viewing the Device Manager tab under Windows 9.x

Attention: Shows were captured from two different computers; that are why they do not have the same devices. If you were to view the same computer through the Device Managers of the two different Windows versions, the lists of devices would be identical.

4.6.1 Troubleshooting of Device Problems

If, a device has installed and working properly, it should be show in the list of Device Manager through their name and icon. Double-clicking it to display its properties should show a message of this device is working properly on the General tab. If you do not see that message, then perhaps you have one of these situations:

Device does not appear on the list. The device is not installed properly, defective, or not PnP-compatible, or your BIOS is not PnP-compatible.

Device appears in the unknown category with a question mark as its icon. You need to run the setup software that came with the device to complete its driver installation.

Device has an X on its icon the device has been disabled. Re-enable it in its Properties box.

Device has an Exclamation Point on its icon there is a resource conflict or a problem with the device. View the status on the General tab of the device's Properties box for details.

The following figure has shown that a device is working properly.

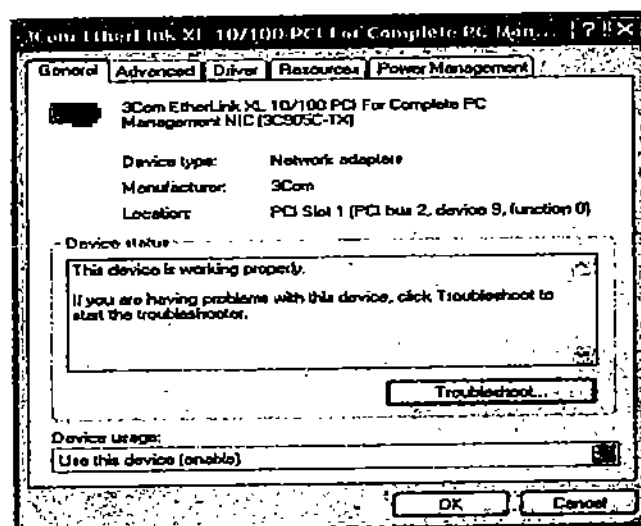


Figure: Showing a device working properly under Windows 9.x

4.6.2 Manual Handling the Device Resource:

Especially on older PCs and in older Windows versions, conflicts may occur between two or more devices that want the same resource. Symptoms of this include the following:

- One or both of the devices are not working.

- One or both of the devices show an exclamation point on its icon in Device Manager.
- For one or both devices, the device status on the General tab of the device's Properties box reports a conflict.
- When you view the device's Resources tab in its Properties box, a conflict shows.

To resolve a resource conflict, change the resource assignment for one of the devices as follows:

1. Open the Properties box for one of the devices, and display the Resources tab. In the Conflicting Device List window, note the other device involved in the conflict.
2. Clear the Use Automatic Settings check box.
3. Open the Settings Based On drop-down list, and choose a different setting. Keep doing this until No Conflicts appears in the Conflicting Device can see in the list of window.
4. Click OK.

Attention: I am showing a device in Windows 98 in following figures, because it is in those lower 9.x based Windows versions that you are most likely to need to do this, Windows 2000 and XP manage resources very well, and so it is unlikely that you will need to do this in one of those versions.

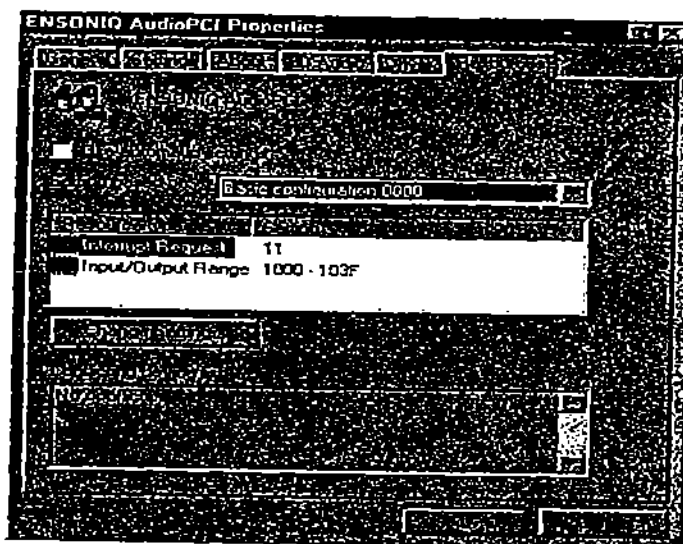


Figure: Viewing the Device has no conflict under Windows 9.x

If none of the choices in step 3 resolves the conflict, you can try manually configuring a particular resource. For example, if the Conflicting Device List window reports that the problem is an I/O address conflict double-click Input / Output Range in the Resource Type list, opening a dialog box where you can choose a different range to use.

What is actually more likely to happen on a modern system, however, is that an error message will appear when you try to change the resource assignments saying that the resource setting cannot be modified? In that case, refer to the Conflicting Device List

window to determine what other device is involved and then try approaching the conflict from that device's properties.

Most new systems today support Advanced Power and Configuration Interface (ACPI) standards. You probably know about ACPI's power management features, but you might not have given much thought to the Configuration portion of that acronym. When ACPI is in effect on a system, it takes over most of the resource assignments. With the result, that the resources for individual devices are not manually adjustable through Device Manager. On the plus side, however, ACPI does a good job managing resources, such that you should not have any conflicts to resolve in the first place.

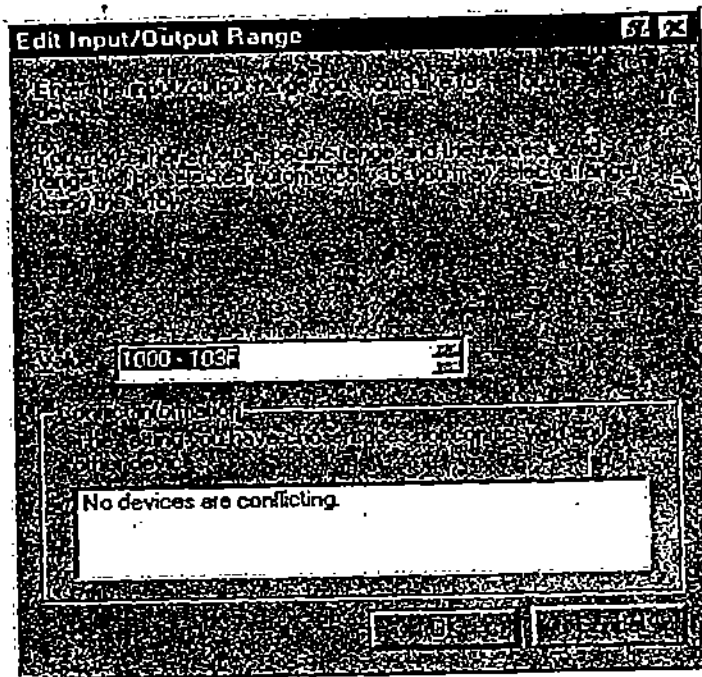


Figure: Manually change a resource assignment if needed.

4.6.3 Removing a Device Driver

After you physically remove a device from the system, its driver might still appear in Device Manager. This could be harmless, or it could cause problem with devices of the same type. For example, if you removed an old modem and installed a new one, the fact that the old modem's driver is still installed can cause problems with some applications recognizing the new modem.

To remove a device driver, delete it from Device Manager:

- Select the device on the list.
- Press Delete.
- Click OK.

Another way to delete a driver (Windows 2000/XP only) is to select it and then click the Remove Device button on the toolbar. Looks like a computer with a red X symbol the one on it.

You can delete the device drivers for hardware that is still physically present, but Windows will redetect the device and reinstall the driver at its first opportunity. One way may possibly to resolve a problem with a device is to remove its driver and then redetect hardware.

4.7 Summary

In this unit, we have discussed about conceptual view of device driver, because the drivers are those, which maintain the accessibility of devices or hardware of computer. We have also explained about kind of device drivers. Further, we have explained how to install these device drivers in the system. Finally, we describe the trouble shooting of device drivers. Further, we have given more explanation about different type of operating system. We have identified the removal various device drivers.

4.8 Model Questions:

Write a short note on the following.

1. What is device driver? Can you explain its importance for computer system?

2. What do you understand by upgradation of new device driver? Describe how to rolling back.

3. What is PnP? How you can install a device when drivers are not available?

NOTES



Uttar Pradesh
Rajarshi Tandon Open University

BLISS-03

**Computer Network and
Security Maintenance**

BLOCK

6

Installation and Maintenance of I/O Devices

Unit 1

Installation and Maintenance of Keyboards and Mice **5**

Unit 2

Installation and Maintenance of Scanners **17**

Unit 3

Installation and Maintenance of CD and DVD Drives **33**

Unit 4

Installation and Maintenance of Floppy Disk Drive **56**

BLOCK INTRODUCTION

Now keyboard, mouse, scanner, floppy disk and CD/DVDs are more common, input and output devices, which are generally used in the computer field. Day to day, some of advancement in the computer field are going to improve the technology. However, as we know the latest type of wireless input and output devices are more popular as compare to wired devices.

Therefore, in this block we will discuss about previously mentioned I/O devices. We will also explain about installation and configuration of these devices. Finally, we consider the description trouble shooting and maintenance of these devices that gives the real boost to generating hardware engineer in the computer field.

We will mainly focus on following point:

Concept of Installation of I/O Device: Concept of installation I/O devices is more necessary, when candidates are wanted to become maintenance engineer in computer field.

Concept of Configuration I/O Device: Concept of configuration of I/O device is also necessary, when candidates are wanted to become maintenance engineer in computer field.

Concept of Maintenance I/O Device: Concept of maintenance I/O device are necessary, when candidates are to hold the knowledge of computer maintenance engineer.

Concept of trouble shooting of I/O Device: Concept of trouble shooting of I/O device is compulsory for candidates of hardware maintenance in computer field.

Unit wise description of block is given as below:

The first unit of this block is concentrated on basics of keyboard and mouse, which create conceptual view of function of keyboard and mouse. We have also included the Installation, configuration and maintenance of keyboard and mouse.

In the second unit of this block, we have devoted on scanners, how it is perform its functions. We further discussed about its installation and maintenance.

In the third unit of this block, we have mainly given the conceptual view about CD/DVD drives, which are also essential for computer professional. In the last, we have explained about configuration and trouble shooting of drives.

In the fourth unit of this block, we have discussed about floppy disk and floppy disk drives and its installation and maintenance.

Suggestion for further reading

1. Foundation of Information Technology, Second Edition, D.S. Yadav
2. Information Technology, Third Edition, Atul Kahate
3. PC Upgrade & Maintenance Guide, Fifteen Edition , Mark Minasi
- 4 Modern All About Motherboard, Manhar Lotia ,Pradeep Nair

UNIT 1: INSTALLATION AND MAINTENANCE OF KEYBOARDS AND MICE

Structure:

- 1.0 Objectives**
- 1.1 Introduction**
- 1.2 Installing the Keyboard or Mouse**
 - 1.2.1 Type of Keyboard and Components**
 - 1.2.2 Type of Keyboard**
 - 1.2.2.1 QWERTY Keyboard**
 - 1.2.2.2 Dvorak Keyboard**
- 1.3 Keyboard Interfaces**
 - 1.3.1 Interfacing Devices**
 - 1.3.2 Wireless Keyboards**
- 1.4 Keyboard Maintenance**
- 1.5 Some Maintenance Tips**
- 1.6 Type of Mouse and Components**
 - 1.6.1 Buttons and Wheels**
 - 1.6.2 Mouse Positioning Methods**
 - 1.6.3 Alternatives to Mice**
- 1.7 Mouse Interfaces**
- 1.8 Mouse and Trackball Cleaning**
- 1.9 Trouble Shooting of Mouse**
- 1.10 Keyboard and Mouse Adjustments in Windows**
- 1.11 Trouble Shooting of Keyboard**
- 1.12 Summary**
- 1.13 Model Questions**

1.0 OBJECTIVES

At end of this unit, student will be able to:

- Understand about the basic function of keyboard and mice.
- Explain the various kinds of keyboard and its components.
- Distinguish in between QWERTY type and Dvorak type of keyboard.
- Know the concept of various types of interfacing devices for keyboard and mice.
- Analyze statistical feature of keyboard and mice.
- Generate awareness about function of mice.
- Identify various trouble shootings keyboard and mice.

1.1 INTRODUCTION

Every personal computer has a keyboard with the universal acceptance of Graphical User Interfaces (GUIs), practically every PC has some kind of pointing device.

The keyboard is the main input device, commonly used by all computers. The keyboard allows for user input action with the computer. Like many typewriters, the keyboard has all the letters of the alphabet, numbers 0 - 9 and additional special operational keys.

When it comes to keyboards, they all have pretty much the same keys, though there are some differences. You have the traditional IBM type keyboard, the newer split ergonomic keyboards, and even wireless keyboards that let you do your typing from your easy chair. They are all the same to troubleshoot.

However, turning to pointing devices, we have tried trackballs, electronic stylus devices, light pens and even a head mouse that tracks the way you are facing. Still, the most popular pointing device seems to be the mouse (either the traditional ball type or the newer optical and wireless). Mice can be more of a pain than keyboards, however, so we will discuss how to handle them in this unit.

1.2 Installing the Keyboard or Mouse

Installing a new keyboard, mouse or trackball are the easiest installations we will encounter. Before you start, the following objects are required.

- Documentation that come with the product to ensure proper installation
- Software or drivers ready to install, if required
- Your new keyboard, mouse, or trackball

To install a keyboard or mouse, follow these steps:

- 1 Turn off your computer if the device you are installing is not a Universal Serial Bus (USB) model.
- 2 Unplug the old device (keyboard or mouse).
- 3 Plug your new device into the appropriate port (keyboard, mouse, or USB) on the back of your computer.

- 4 Turn your computer on (if it is off).
- 5 Windows will detect almost all keyboards and mice and allow them to function at a basic level. If desired, install the software that came with the device to add additional functionality.

1.2.1 Type of Keyboard and Components

With all of its moving parts, the keyboard has many potential sources of problems. Understanding those sources requires understanding the types of keyboards and their parts; here is a look at them.

1.2.2 Type of Keyboard

Carpal Tunnel Syndrome may occur as a user does repetitive motion on his or her keyboard. There have been several methods to help prevent this from happening. One of the most common ways in preventing this is purchasing a Natural Style ergonomic keyboard. This type of keyboard allows the hands to be in a more suitable position. These keyboards can be purchased from any computer retailer. It is recommended you test the keyboard before purchasing it. The keyboard can be difficult to use for some users. Below is an illustration of what type of keyboard may look like.

Natural Keyboard



1.2.2.1 QWERTY Keyboard

Developed by Sholes in 1868s and it is now the official standard of computer keyboards. The layout of the QWERTY keyboard was designed for two finger typing and was designed to help prevent stuck keys. Now with new technology, stuck keys are no longer an issue; and with the position of the keys, the hands are held close together helping to cause Carpal Tunnel Syndrome. Given all the above problems, this is the most widely used keyboard because of the lack of options, can be used without training, is used because of the difficulty of having to teach a different keyboard, and is cheap and easy to find.

QWERTY KEYBOARD

~	!	@	#	\$	%	^	&	*	()	-	+	Delete
Tab	Q	W	E	R	T	Y	U	I	O	P	{	}	\
Caps	A	S	D	F	G	H	J	K	L	:	"	'	Enter
Shift	Z	X	C	V	B	N	M	<	>	?	/		Shift
Ctrl		Alt									Alt		Ctrl

1.2.2.2 Dvorak Keyboard

Dr. August Dvorak developed it in 1936 as a solution to the QWERTY keyboard. The layout of the keys on this keyboard makes it easier and faster to type on when compared with other keyboards.

DVORAK KEYBOARD

	1	2	3	4	5	6	7	8	9	0	-	=	Delete
Tab	?	<	>	P	Y	F	G	C	R	L	{	}	
	A	O	E	U	I	D	H	T	N	S	~	.	Enter
Shift	:	Q	J	K	X	B	M	W	V	Z			Shift
Ctrl		Alt									Alt		Ctrl

There is also other number of ways to distinguish one keyboard from another, but the most obvious way is the number and arrangement of keys.

You will first look at a couple of keyboards you are not likely to see anymore. The XT keyboard and the AT keyboard were introduced with the IBM PC and the IBM PC-AT, respectively.

The XT keyboard had 83 keys, and the function keys and numeric keypad were at the right of the other keys. There was no separate, set of arrow keys; you had to press Num Lock to enable the numeric keypad as an arrow pad.

Improvements of the AT keyboard over the XT keyboard include a larger Enter key and the addition of a SysReq key intended for use with OS/2 (IBM's attempt at an operating system of that era). It also has 10 function keys either along the top or on the left side of the keyboard (labeled F1-F10).

The main difference inside the XT and AT keyboards is that the XT keyboard puts the keyboard microprocessor in the keyboard, and the AT keyboard assumes that the keyboard microprocessor is on the system board. They are generally incompatible: you cannot use an XT keyboard on an AT, or vice versa. Clone keyboards generally get around this by putting an XT / AT switch on the keyboard.

Both XT and AT keyboards are long obsolete, replaced in the marketplace by the enhanced keyboard. It is compatible with all other keyboard interfaces (AT and XT), but you might need to use an adapter plug to change the pins/size of the connector to make it fit. The enhanced keyboard layout has at least 101 keys. The extra keys come from adding two more function keys (F11 and F12) and adding a separate arrow keypad.

Windows keyboards add three keys to the standard 101-key keyboard, typically on the same row as the spacebar. The new keys are two Windows keys (one on either side of the spacebar) that open the Windows Start menu and an Application key that displays the same context menus you get when you right-click an item.

Attention: Many newer keyboards have additional keys that can speed up your productivity. Some keyboards provide a large number of specialized keys that help you connect to, browse the Internet and even operate your CD and multimedia applications. Some keys through the help of drivers and special software can configure to perform

your favorite tasks or to work with programs of your choosing; these are called as not keys.

The broken keyboard designs (also called ergonomic keyboards), such as the Microsoft Natural Keyboard look like a 101-key keyboard to the PC hardware. Because they are compatible with the existing hardware, they are easy to install, and although they take a little getting used to, they are easier on the wrists than standard keyboards. People who do a lot of typing might prefer these.

1.3 Keyboard Interfaces

1.3.1 Interfacing Devices

A keyboard can have three possible interfaces:

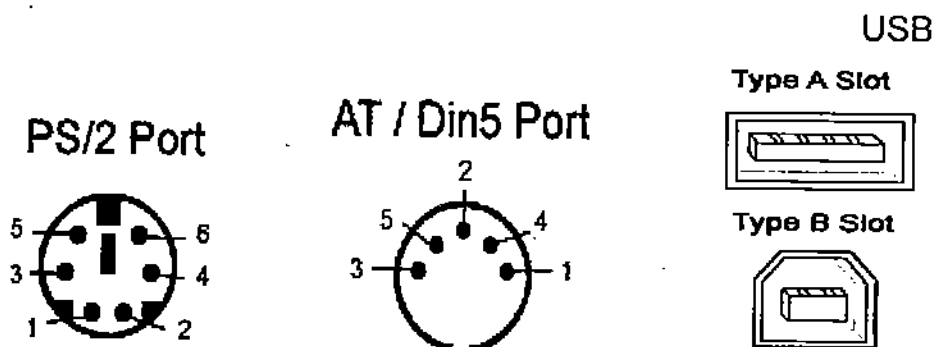
- AT (DIN)
- PS/2 (mini-DIN)
- USB

The motherboard will have either PS/2 or AT-style keyboard connector built into it. (That is the case as of this writing, but manufacturers will soon be phasing out the inclusion of any keyboard interface because they assume consumers will be switching to USB keyboards). The keyboard interface depends on the motherboard form factor; AT motherboards have the AT interface, and ATX motherboards have the PS/2.

Another name for the AT keyboard connector is DIN. It is a large round plug with five big pins.

The PS/2 style is also called mini-DIN, and it has six small pins; in recent years, USB has become the interface of choice for keyboards (and mice). The Microsoft Natural Keyboard and several Microsoft mice are examples of devices that are USB in design, but some versions may include a USB-to-PS/2 adapter with the product. USB devices have several benefits, which mostly derive from USB being a newer, faster technology. A USB connection is faster than the other, older I/O ports on your computer, such as COM and parallel ports. And, USB device does not require any special device drivers to operate it: (A driver is required, but USB drivers cover a wide range of devices, so for example, a USB pointing device driver covers mice, touch pads and writing boards). You simply plug the device in and it works (ideally that is) USB devices also require no conflict resolution for competing resources such as COM ports, which require Interrupt Requests (IRQs) and other precious system resources to operate.

About the only drawback to using a USB device is backward compatibility. Although Windows 98 and higher are USB aware, Windows 95 and NT4 are not and therefore, incompatible with USB devices.



1.3.1 Wireless Keyboard and Mouse

A wireless keyboard or mouse frees you of the connection between the PC and the keyboard, so you can sit back in your recliner and type on your large screen monitor from across the room.

They work like this: you get a receiver with the wireless keyboard that plugs into your computer, usually through the keyboard or USB port. This cabled receiver then sits on top of your monitor (in the case of infrared systems) or anywhere within 30 feet or so of your keyboard or mouse for Radio Frequency (RF) models.

Some of these devices use Infrared (IR) frequencies to transmit data, meaning that they have high bandwidth (not really an issue with something that transmits as little data as a keyboard or mouse does) and are immune to outside interference. Use a cordless telephone or other infrared devices (such as a TV remote control) in the same room and it will have no effect at all. For the infrared frequencies to communicate, however, the receiver (on your PC) and transmitter (on your mouse or keyboard) have to be within the line of sight.

Wireless keyboards are more common radio controlled keyboards. Newer wireless devices operate in the 900-MHz frequency band and have a range of about 30 feet. Because RF signals pass through most objects, your mouse and keyboard, do not have to be within the line of sight of the receiver, a big advantage over IR devices.

If you are using an RF device, there is a slight risk of interference with other wireless devices (such as cordless phones) that also use the 900MHz band. Practically, however, this is not much of an issue; if you do experience interference, just put a little distance between the two devices; and the problem will probably fix itself.

Generally, wireless devices are more expensive than wired devices and input devices are no exception. Although one manufacturer's advertising and magnifies wireless keyboards as a wonderful way to reduce cabling needs, the lack of one or two cables does not make up the price difference. You can find the cheapest keyboards for under \$10, with wireless models starting at around \$40.

Personally, I have grown to like wireless input devices. I have been using a Logitech wireless keyboard and mouse for about a year now, and I like being able to lean back in my chair, put my feet up on my desk and type away with my keyboard in my lap, not worrying at all about whether the cord will reach. If you do not mind replacing batteries every few months (and do not mind the higher initial cost), doing away with the cords is a good thing.

1.5 Keyboard Maintenance

The major aspect of maintenance for keyboards is approximately nil. That is from spilling things into the keyboard. Protecting keyboards with plastic covers are necessary step to keep keyboard safe. This is one approach and another is just to be careful.

Some people like to clean their keyboard by pulling the key tops off and cleaning them individually, plus cleaning underneath them. Do it if you want (but make sure you can put them back in the right places), but personally I do not do this. I clean my keyboard by doing the following:

1. Turn the PC off, and disconnect the keyboard.
2. Turn it upside down and shake it so that all the debris falls out onto the floor. (What I did not say "Put down a towel to shake it over?" I should have).

3. Use a can of compressed air to blowout any remaining particles under the keys.
4. Clean the key tops with a paper towel dampened with a cleaning solution for computers. If you have to remove some keys to get them cleaned, do so, but personally I try not to do so.

In extremis, you can soak a sticky keyboard in water overnight or even run it through your dishwasher, but make sure it is completely dry before you using it. I would not put water on a keyboard unless it was gross, though like if it had spilled a can of sugared soda on it.

What to Do When a Keyboard Fails: Because most keyboards are so inexpensive, (with the exception of some ergonomic and wireless keyboards) you may want to view the keyboard as disposable. You can do some simple things before throwing out a keyboard.

Make sure about plug in: On the back of computers that use PS/2-type, connectors are two identical ports: the mouse port and the keyboard port. Make sure you have plugged the keyboard into the correct port. You will get an error when turning on the computer if the mouse or keyboard is connected the wrong port.

Make Sure about BIOS and Windows: If there is a keyboard failure, you should see a message when your computer boots up. A PS/2 style or AT-style keyboard will detected as a legacy keyboard. Pay close attention and if you see the message, try checking the connection first before you move on to more drastic measures.

If the BIOS does not display any errors on screen (such as Keyboard not present or Keyboard failure) when you boot up, make sure Windows recognizes your keyboard by checking Device Manager. You can access Device Manager in Windows by opening the Control Panel and double clicking the System icon. Select the Device Manager tab in the System Properties dialog box. Your keyboard is list in the Keyboard category. If it is in the list, make sure it is the proper make and model.

To access Device Manager in Windows 2000, open the Control Panel and double-click Administrative Tools. Next, double-click Computer Management, and expand the System Tools category to find Device Manager. Your keyboard should list in the Keyboards category.

Note: If a USB keyboard does not work at a command prompt (prior to Windows 2000, check to see whether there is a Legacy USB setting in Basic Input Output System (BIOS) Setup that you can enable. This will help the system recognize USB input devices outside of Windows.

Fix Bent Pins: Check your PS/2-type connector that plugs into the motherboard for any bent or missing pins. This is often happens if someone (usually children, but adults can do this too if they are not paying attention) tries to force the connection in the wrong orientation. What happens is that the pins that are not line up with the correct holes on the motherboard connector bent back as you force the connection. There is simply no place for them to go but sideways. If your pins are bent, carefully straighten them with a pair of slim, needle-nosed pliers. Be careful not to break off the pins as you straighten them.

Connect It to another Computer: If you are fortunate enough to have more than one computer in your home (most businesses have several), you can try swapping keyboards with a machine that you know is not suffering keyboard problems. If the keyboard that you suspect is having trouble does not work on the "good" computer, you know that keyboard is bad. If the new keyboard does not work on the machine you are

testing, you may have a faulty keyboard connection on the motherboard. Unfortunately, there is not much you can do about this yourself. If your computer or motherboard is under warranty, or you have a maintenance agreement, you might be able to send your computer to the manufacturer and let them replace the keyboard connection. If you are doing the work yourself, you may have to replace the motherboard.

Attention: A good workaround for a faulty keyboard connector is to purchase a low-cost PCI card with a built-in key card connector; this is a good way to get a new keyboard connector without going to the cost and trouble of replacing the entire motherboard. (Similarly, you can just buy a new USB keyboard and connect it to your PC's USB port bypassing the keyboard connection completely).

Check the Keys: If only one key is malfunctioning, check the spring under the key to see whether it springs up and down, as a key should. Remove the key by grabbing it with your fingers and pulling up. For the tough keys, fashion a hook from a paper clip or, again, use a chip puller. You will see a spring under the key. Replace the key top, and see whether the problem goes away. If not, try pulling the spring out just a little. Then replace the key top.

Some keyboards use rubber cups instead of springs. The way to, either the cup or spring is design to keep the keys from being on all the time.

Test Pin Voltages: The voltage between pin-4 and each of the other pins should be in the range of 2 to 5.5 volts Direct Current (DC). If any of these voltages are wrong, the problem probably lies in the PC, the system board in particular. If they are okay, the problem is probably in the keyboard. Note that the pins are not number consecutively for either a DIN or a mini-DIN.

If you hold a DIN so that the pins are at the top and count from left to right, the pins are numbered 1, 4, 2, 5, and 3; the 2 pin is at the top center point. Holding a mini-DIN with the single groove in the connector at the top and starting from the bottom left and moving clockwise, the pins are numbered 1, 3, 5, 6, 4, and 2; the groove is between pins 5 and 6.

Check the Cable Continuity: Next, test the continuity of the cable. Turn the keyboard upside down, so that the cable is coming out of the back of the keyboard to the right. Remove the two screws: The bottom plate will swing back and up for removal.

You will now see that the cable splits to a single wire, which is ground to the bottom plate. You will also see a cable with a flat-jaw connector. Push apart the jaws of the connector to release. You can then use an ohmmeter to test each of the five wires for continuity. An ohmmeter is a device that electricians (and computer technicians) use to test electrical circuits for power and detect short-circuit.

Replace the Keyboard: If you have gotten this far and have had no luck, do not despair. Many keyboard problems cannot fix, and the things are not that expensive anyway. As you are buying for a new keyboard, consider these factors:

- How it feels when you type on it :(in store displays will help)
- Connector type
- Wired or wireless
- Special keys such as volume control, shortcuts to e-mail and so on
- Split keyboard (ergonomic) or traditional

1.6 Types of Mouse and Components

Mice are in some ways simpler devices than keyboards; after all, most have only two or three buttons and maybe a wheel as compared to the standard keyboard's 101 keys. (Some new mice do have up to six buttons.) However, most mice have some moving parts that can go seriously wrong.

1.6.1 Buttons and Wheels

All mice for PCs have at least two buttons: left and right. (You may occasionally find a very old Mac mouse with a single button.) You cannot go wrong choosing a two-button mouse, but a number of more modern mice have extra buttons and wheels to make scrolling, zooming, and selecting much easier. Microsoft's Intel Mouse Explorer has five programmable buttons that can use with standard programs and Internet-related applications; it also has a small wheel (located between the left and right mouse buttons) that zooms in and out or scrolls.

1.6.2 Mouse Positioning Methods

Mice work in this way: you move them, they figure out somehow, how much you have moved them, and then they transmit that information to the computer.

The older type of mice is mechanical meaning that they use some mechanical method for detecting motion. In most mice, the primary mechanical part is a ball on the bottom of the mouse. Take the mouse apart (remove the retaining ring, turn the mouse over, and the ball falls out, generally rolling under something), and you will see three little wheels that turn when the ball moves against them.

Two of those wheels are monitored electronically; when they turn, they transmit to the computer how much they turned. The two wheels are perpendicular to each other, so one tracks X-axis motion and one tracks Y-axis motion. The third wheel just balances the first two.

An alternative kind of positioning mechanism is in the optical mouse, sold by Microsoft, Logitech, and others. This type of mouse has no moving parts. The mouse has optical sensors instead of a physical mouse ball, and can be used on any surface.

Note: Do not confuse an optical mouse with a wireless mouse. A mouse can be one or the other, neither, or both. An optical mouse is one that lacks built-in uses optical sensors for tracking. A wireless mouse is one that lacks a cord and uses either IR or RF signal to communicate with the PC.

Optical mice operate by using a red Light Emitting Diode (LED) that takes 1500 snapshots per second of the surface below it. By comparing the images dynamically, they can determine speed, direction, and distance. You can use these mice on top of just about any surface that is relatively flat. However, you will want to avoid using these optical mice on surfaces with solid color or on glass or reflective surfaces. You will also want to avoid the patterns, which are highly repetitive.

Note: The new Microsoft optical mice are a boon to game players. In addition to having precise positioning, they are all USB mice. Normal mice operate at 40MHz (the frequency their position is updated by Windows) because of the interface, but USB mice operate at 120MHz. With frame rates if some games shooting at more than 40MHz, having a mouse that can keep up with your game is necessary.

1.6.3 Alternatives to Mice

Trackballs are pointing devices that have the ball (normally located underneath and inside of a mouse) on the top of the unit. Instead of moving a mouse around, you move

the trackball with your finger. It is just like the trackballs on some ideally older machines or like using an upside down mouse, if you want to think of it that way.

Many people prefer trackballs to standard mice because you have fine control over the position of the cursor and do not have to move anything around on your desktop. Trackballs are ideal for certain professions, such as architects and graphic artists, who need a high degree of pointing precision as they work. Trackballs are also an attractive alternative for young children and people with certain disabilities because they often find them easier to use than mice. Some laptops even use this technology as their primary pointing device.

Another mouse alternative is a touchpad, popular on notebook computers. A touchpad is a touch sensitive rectangular pad. You move your finger over the pad to move the pointer on screen.

1.7 Mouse Interfaces

A mouse can connect to the PC using USB, PS/2 legacy serial port, or a proprietary bus interface. That list is in order of modernity, from newest to oldest. You will probably never encounter a proprietary bus interface mouse and very seldom a legacy serial one.

1.8 Mouse and Trackball Cleaning

On a mechanic mouse or trackball, the little wheels get particles stuck to them or hair wrapped around them, so you must clean them. Just remove the mouse ball and examine the wheels. I have used rubbing alcohol, a toothpick, and a cotton swab to get the particles off the wheels. You may also need to clean the ball with a bit of alcohol. Clean the mouse about twice a year or more if your desk is covered with dirt or dust, which is just gunk in its fetal stages.

On an optical mouse, all you have to do is clean the sensors; you can often just polish them with soft cloth.

1.9 Trouble Shootings of Mouse

You can troubleshoot a mouse by following these steps:

1. Make sure your mouse is plugged in securely.
2. Check the driver. Is the mouse driver is setup correctly? Is it there in the first place?
3. Clean the mouse, and shake any debris off the mouse pad.
4. Check the interface at the end of your mouse cable and at the motherboard to see whether it is USB, PS/2 or the serial port. See if there are any obvious problems such as bent pins (for PS/2 and serial port mice only) that can be straightened.
5. Ensure that Windows recognizes the mouse and that it has listed in Device Manager as properly installed.

If all this fails, and your computer is not the culprit, you will need to replace your mouse. Fortunately, mice are quite affordable; standard mice run around \$10, with the fancy optical models starting, at around \$20.

1.10 Keyboard and Mouse Adjustments in Windows

When a keyboard or mouse is not working optimally, you can sometimes get relief by adjusting its settings in Windows through the Control Panel.

Keyboard settings you can change in the Keyboard applet include these:

- **Repeat Delay:** This is the amount of time between holding down a key and having it start to repeat rapidly in typing.
- **Repeat Rate:** A key repeats after the delay time has passed at this speed.
- **Cursor Blink Rate:** In applications that accept text, you may see a blinking vertical cursor showing where the text will appear; this setting controls the blink speed.

Mouse settings you can change in the Mouse applet include the following:

- **Pointer Speed:** This is the amount that the pointer moves across the screen in relation to the amount you move the mouse. Sometimes this is called pointer sensitivity.
- **Pointer Scheme:** This controls the look of the various mouse pointers. You might switch to a large pointer set for someone with limited vision, for example.
- **Switch Primary and Secondary Buttons:** Left handed users may prefer to use the mouse with the left hand and to switch button functionality so that the primary button is under the index finger.
- **Double-Click Speed:** This controls how close together the two clicks in a double click need to be perceived as a double click rather than two separate clicks. A beginner having trouble double-clicking might benefit from having this setting turned down.

It is important to note that the options described pertain to the default drivers for keyboards and mice in Windows. If you install the special software that comes with your keyboard or mouse, the Properties box for Keyboard or Mouse in the Control Panel may display very different options and probably will offer more sophisticated options and settings.

1.11 Trouble Shooting of Keyboard

The following is a summary of troubleshooting tips for your keyboard:

- Make sure that the keyboard is plug in and the connection is secure.
- Ensure that there are no BIOS warnings at boot up and that Windows recognizes the keyboard. Check the cable connector for any bent or missing pins.
- Swap the keyboard you suspect is malfunctioning with one that you are sure works to see if the problem lies with the computer and not the keyboard.
- Check the keys for proper functionality. They should all spring back to their original state when you press and release them.
- Test pin voltages on the keyboard cable.
- Check for cable continuity within the keyboard housing.
- Disassemble the keyboard and check all internal parts.
- Replace the keyboard if none of the previous tips work and your computer is not at fault.

In addition, here are a few things to try when your mouse is not working correctly:

- Ensure that the mouse is plug in the security.

- Check the mouse driver in Windows to make sure it is correctly install.
- Clean your mouse to remove built-up particles.
- Check for bent pins on the cable.
- Replace your mouse if nothing else works.

1.12 Summary

In this unit, we have discussed about some important input and output devices like keyboard and mouse. We have also described installation and configuration of these input devices. We have introduced different components of keyboard and mouse. Finally, we have explained about trouble shootings and maintenance of keyboard and mouse.

1.13 Model Question:

Write a short note on the following.

1. What are roles of keyboard? Can you explain its components?

2. What role has been played by mouse in the computer? Explain the different component of mouse.

3. Write the comparative picture of keyboard and mouse.

UNIT-2 : INSTALLATION AND MAINTENANCE OF SCANNERS

Structure:

- 2.0 Objectives**
- 2.1 Introduction**
- 2.2 Connecting a Flatbed Scanner**
- 2.3 Types of Scanners**
 - 2.3.1 Flatbed scanners**
 - 2.3.2 Sheet fed Scanners**
 - 2.3.3 All in one scanner**
 - 2.3.4 Film scanners**
 - 2.3.5 Drum scanners**
- 2.4 Function of Scanner**
- 2.5 Scanner and Computer Interfacing**
 - 2.5.1 USB Connector**
 - 2.5.2 Parallel Connector**
 - 2.5.3 SCSI Connector**
- 2.6 Image Quality Measurement**
- 2.7 Picture of Scanner in Windows**
 - 2.7.1 TWAIN Software**
 - 2.7.2 Windows Image Acquisition (WIA)**
 - 2.7.3 Optimization of Size and Quality of Scan Image**
 - 2.7.4 Selection of Appropriate Image Resolution**
 - 2.7.5 Selection of Appropriate File**
 - 2.7.6 Selection of Appropriate Image Mode**
 - 2.7.7 Optical Character Recognition**
- 2.8 Scanner Maintenance**
- 2.9 More Tips on Troubleshooting**
- 2.10 Summary**

2.0 OBJECTIVES

At end of this unit, student will be able to:

- Understand about the basic function of scanner.
- Explain the various kinds of scanner and its components.
- Distinguish in between flatbed type and sheet fed scanners.
- Know the concept of various types of interfacing devices for scanner.
- Analyse statistics of picture under the Windows environment.
- Generate awareness about all in one machine perform operation like scanner.
- Identify various trouble shootings of scanners.

2.2 Connecting a Flatbed Scanner

The following are the basic steps for connecting the most popular type of scanner the flatbed scanner. For in depth coverage on how to troubleshoot scanner related problems, refer to the later sections of this unit.

Before you start, you may need to include the following tools:

- Scanner documentation or user manual required.
- Scanner software or drivers ready to install, if required.
- A cable for connecting the scanner to the computer is required.
- New flatbed scanner, which may required for connecting.

To connect a flatbed scanner, follow these steps:

1. If you are installing via a parallel or Small Computer System Interface (SCSI) port, turn off your computer. If you are installing via a Universal Serial Bus (USB) port, leave your computer on.
2. Plug your new scanner into a power source and then connect it to the appropriate port (USB, parallel or SCSI) on the back of your computer. If a scanner shares a parallel port with the printer, power off the printer and connect the printer and the scanner to your PC, as directed in the scanner's instructions. If you are connecting via SCSI, you will probably need to install a separate SCSI card in your computer and connect the scanner to that board.
3. Turn your computer on, and then turn on your scanner.
4. Windows should recognize the new scanner and install the appropriate drivers.
5. You may be prompted to insert the scanner's installation disk or CD at this step.
6. If your scanner came with its own software (most scanners typically come with some graphics program or another), install that software now.

Attention: Some scanners have a lock to protect the Charge Coupled Device (CCD) from being damaged during shipping. If your unit has such a lock, do not

forget to unlock the CCD before you first use the scanner. It is probably a switch on the bottom.

Attention: Some high-end scanners might require a manual calibration on installation. Scanning a special shaded card that comes with the scanner and then configuring various settings on the software of scanner control panel typically do this.

A scanner works a little like a traditional office photocopier except that the result is a digital computer file, not a printed copy.

2.3 Types of Scanner

You can install several types of scanners on your system: There is little or no difference in how each interacts with the computer; each uses one of the standard interface types (USB, parallel, or SCSI) and each is recognized as a scanner in Windows. However, there is a difference in their physical dimensions and arrangement and operation of the physical parts inside:

2.3.1 Flatbed scanners

The most popular type of scanner for home use is the flatbed, or desktop, scanner. They make it easy to scan papers, books, and any other item that you can lay flat between a glass bed (plate) and the scanner's top cover. The image is scanned via a scan head that moves across the face of the original document.

2.3.2 Sheet fed Scanners

Sheet-fed scanners are like flatbed scanners except that the scan head are fixed and the original document moves across the head like in a fax machine. Whereas flatbed scanners can scan just about any item, that can fit on the glass plate, including three-dimensional objects. Sheet fed scanners can scan only flat pieces of paper. Sheet-fed scanners were popular in the late 1990s but today are rarely seen.

2.3.3 All in One scanner

A popular option: in home offices and small offices is the "All-in-One" machine that scans prints, faxes, and copies. These units, popularized by Hewlett-Packard (HP), effectively merge a flatbed or sheet fed scanner with an inkjet or laser printer and a fax machine. Such a unit may have its own proprietary integrated software that drives it through Windows rather than relying solely on separate Windows drivers for each component.

In the early days, the scanners are inexpensive models were handled. They were somewhat like the bar code readers you may see in retail stores, but wider (about 5 inches across, usually). You did drag them across a page to scan it, a half page at a time, and then software would knit the two halves together. They never worked all that well and are now obsolete. In the following figure, we have shown the All in One Machine which can scan, print and copy the document.

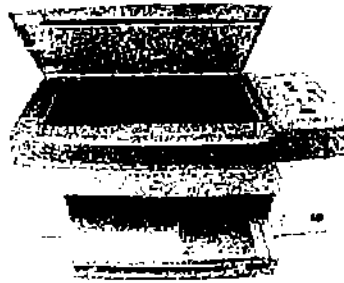


Figure: All in One Machine

2.3.4 Film scanners

There are specialties of scanners available that scan from 35mm negatives or 35mm slides. Because the originals are so small, such scanners must scan at a high resolution to produce an acceptable quality output, usually 3000 to 4000 dots per inch (dpi). One of these is an indispensable piece of equipment for many professional graphic design businesses and magazine publishers, but the average consumer would look elsewhere.

2.3.5 Drum scanners

If you want high quality black and white or color scans, such as the kinds of scanners required by the magazine, newspaper, and book publishing industries, you need to go all the way up to an expensive drum scanner. This type of scanner mounts the original document on a rotating glass cylinder, called drum. At the center of the cylinder, the sensor splits light bounded off the document into three beams. Each beam is then sent through a color filter into a photo multiplier tube where the light is changed into an electrical signal. Drum scanners are more expensive than consumer quality flatbed scanners and they typically are connected to a computer system via a SCSI.

In this unit, we will spend more time on the flatbed scanners as compared to other technologies because they are the most prevalent in the average technician's work environment. Just be aware that the other types do exist, in case you should encounter one at some point.

2.4 Function of Scanner

Most flatbed scanners are composed of the following parts:

- Glass bed (or plate), on which the source document is placed facedown.
- Lamp, used to illuminate the source document.
- Mirrors, used to reflect the image of the source document.
- Filters, used to adjust the image of the source document.
- Lens, used to focus the image of the source document on the CCD array.
- CCD array, used to turn reflected light into an electrical charge.

- Scan head, which contains the CCD array, mirrors, lens and filter.
- Stabilizer bar, to which the scan head is attached.
- Belt, attached to the stepper motor and used to advance the stabilizer bar.
- Stepper motor, used to drive the stabilizer bar.
- Cover, used to provide a uniform background for the scanned document and to keep you from being blinded by the scanner lamp.

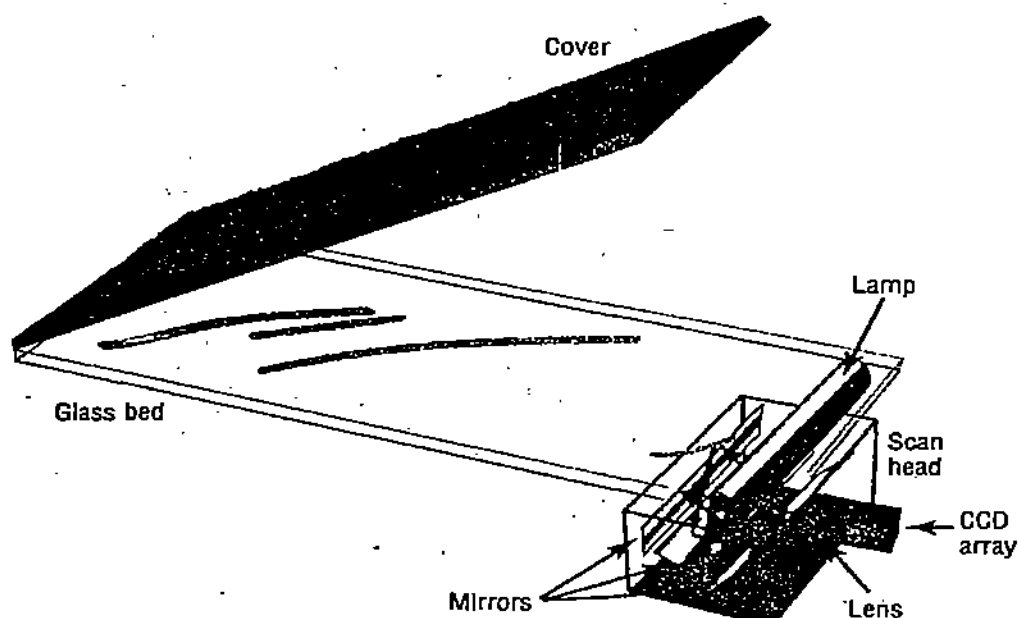


Figure: The major parts of a flatbed scanner

You prepare for a scan by placing the source document facedown on a glass plate. You then close the scanner's cover, which provides a uniform background that the scanner software can use as a reference point for determining the size of the scanned document.

When you press the button to start the scan, the lamplights to illuminate the source document and the stabilizer bar rolls the document from one end to the other. As the scan head, which is attached to the stabilizer bar travels across the face of the document (in what is called a pass), light is reflected off the document through a series of mirrors, filters, lenses and then onto the CCD array. Some scanners will do a draft pass first to send preliminary info into the scanner software that can be used to indicate the image area if desired.

Attention: Older scanners used a standard fluorescent lamp, but most new scanners use either a Cold Cathode Fluorescent Lamp (CCFL) or xenon lamp.

The CCD is actually collection of light-sensitive diodes, called photo sites. The photo sites convert the reflected light into an electrical charge; because each photo site is sensitive to slight variations in light, the brighter the light that hits a photo site, the greater the electrical charge generated.

Most consumer-grade scanners use a single pass to scan the original document, but some higher end models use a three-pass method. In the single-

pass method, the lens splits the image into three identical versions of the original. These images are, then passed through three-color filters (red, green and blue) to separate sections of the CCD and combined to create a single full-color image. In the three-pass method, each pass of the scan head uses a different color filter (red, green, or blue) between the lens and the CCD array; assembling the three filtered images results in a single full-color image.

Some inexpensive flatbed scanners use a Contact Image Sensor (CIS) instead of a CCD array. The CIS replaces the entire CCD/mirror/filter/lens/lamp mechanism with rows of red, green, and blue Light Emitting Diodes (LEDs). The image sensor is placed very close to the glass plate, and the LEDs combine to provide a bright white light. The same sensors then capture the illuminated image.

Attention: Although CIS scanners are small, lighter and less expensive than similar CCD scanners, they do not deliver the same image quality as their higher priced CCD. Look in the scanner's specifications before you buy to see if it is CIS or CCD.

2.5 Scanner and Computer Interfacing

Three main interfaces connect scanners to personal computers: USB, parallel (on older models only) and SCSI:

2.5.1 USB Connector:

The USB interface is the easiest way to connect a consumer-grade flatbed scanner. Just connect a cable from your scanner's USB output to a USB input on your PC. Because USB is a "hot" interface, you do not have to turn off your computer to make your connection and Windows should recognize the new device as soon as it is plugged in and turned on. A scanner can support either USB 1.1 or USB 2.0 as its standard. If connected to a PC that also supports USB 2.0, there is a big speed difference in the data transfer with a 2.0 scanner interface. However, the mechanical parts of the scanner are also a speed bottleneck usually more so than the USB connection.

2.5.2 Parallel Connector:

Before USB became popular, parallel was the interface of choice for consumer-grade scanners of all types. It has disadvantages over USB. Because the parallel port may also be supporting a printer, the two devices must share it. This is achieved with a pass through or Y-splitter that typically comes with the scanner. This type of connection can be problematic, especially if you try to use the scanner and printer at the same time.

2.5.3 SCSI Connector:

The faster SCSI interface is used for the high data transfer rates inherent with drum scanners. Most drum scanners include a dedicated SCSI card you have to install in your computer, but many also let you use a standard SCSI controller. Some high-end flatbed scanners are also SCSI.

If you plan to buy a typical inexpensive flatbed scanner and you have a choice, go with a USB model. They are the easiest to hook up and do not interfere with other peripherals in your system (such as your printer).

2.6 Image Quality Measurement

Even the lowest-quality scanners today deliver a resolution of at least 1200 x 1200dpi. The highest quality scanners today can exceed 4800 x 4800 dpi. The dpi is determined by the number of sensors in a single row of the CCD or CIS array (which determines the x-direction sampling rate) and by the precision of the stepper motor (which determines the y-direction sampling rate).

For example, a scanner with 10,200 sensors in each horizontal row delivers an x-direction sampling rate of 1200. (10,200 sensors divided by 8.5 inches = 1200 sensors/inch). If the stepper motor can move the stabilizer bar in increments of 1/ 1200th of an inch, the Y-direction sampling rate is also 1200 resulting in a 1200 x 1200 dpi resolution.

The scanning software used by the scanner can artificially enhance resolution. Some software programs interpolate extra pixels between the actual pixels, thus increasing the apparent resolution. For example, software that puts one extra pixel between each real pixel turns a 1200 x 1200dpi scanner into a virtual 2400 x 1200 scanner. Interpolation works by averaging the values of two adjacent pixels and inserting an extra pixel between them that is half way between the two. When a scanner's resolution is advertised as optical resolution, that number is the actual number of sensors on the CCD. Optical resolution is real resolution. Resolution advertised as enhanced or digital is interpolated.

The other factor in image quality is the sharpness of the image. Sharpness is determined primarily by the quality of the lens optics and the brightness of the light source. The image also depends upon the good quality of the lens, the best brightness of light and sharpness of scanning process.

Color fidelity is measured in terms of bit depth. Almost all color scanners deliver 36-bit color, and some higher priced models promise 40 or 48-bit color. Although these higher-bit scanners process colors with the higher bit depth, they still output in 24-bit color, so you might not notice any measurable difference in the color of your scans. The extra bits are used for on the fly color correction. The scanner selects the best 24 bits it gets from the scan and sends them on to the computer; the other bits are discarded.

When scanning in grayscale mode, no scanner on the market today can produce more than 256 levels, even if it boasts up to 48-bit color scanning. To understand why, you need to realize that color scanning outputs in 24-bit color with 8 bits of red, 8 bits of green, and 8 bits of blue -in other words, 256 levels of each of those colors. With grayscale, you have only one color (black), so you are limited to 8-bit output for it (256 levels).

2.7 Picture of Scanner in Windows

How you interact with the scanner through Windows depends largely upon whether your version of Windows directly supports your scanner.

Windows Me and Windows XP support a technology called Windows Image Acquisition (WIA) for some scanner models. If you have one of these Windows versions and your scanner is compatible with WIA for that version, you do not have to install the software that comes with the scanner. You just plug in, go, and use the Windows Image Acquisition (WIA) through the Scanner and Camera Wizard to do your scanning. If there is not direct support, you must install the scanning software for the scanner and interact with it through some scanning application (usually third-party) via the TWAIN driver.

Especially with Windows Me or XP, there have been some glitches noted where Windows detects the scanner automatically and appear to be installing a driver for it, but then the scanner does not work through Windows Scanner and Camera Wizard interface. This spotlights an important fact: Plug and Play does not equal WIA necessarily. Even though Windows maybe able to detect a scanner, it does not necessarily mean that, the scanner will work with WIA.

This business of having to use the scanner's own software is really no big deal. That is how it has always been throughout the history of scanning, before WIA was invented. So do not panic if you have a scanner that is not on the Hardware Compatibility List; just download the latest drivers for it for your operating system from the manufacturer's Web site, and it should work just fine.

2.7.1 TWAIN Software

There are hundreds (thousands, even) of new and old scanner models there and each, one is a little different. So how is it that they can all communicate with Windows applications? In a word: TWAIN.

TWAIN is a universal software interface driver that acts as an interpreter between the scanner and any TWAIN-compliant applications on your computer, such as a graphics program with a scanning capability. When you install the software that comes with your scanner, you install the TWAIN driver and a TWAIN-compliant graphics program and the two of them work together to enable you to scan.

Attention: TWAIN is that rare computer term that is not an acronym. (It actually comes from the phrase never the twain shall meet because the driver sits between the software and the scanner) For those if you who are acronym inclined, however, feel to use the following pseudo-acronym for TWAIN: Technology without an interesting name.

If you have Windows XP or Me and your scanner is on the Hardware Compatibility List for your operating system version, you do not need TWAIN because Windows supports the scanner directly through WIA. However, you are still free to install the scanner's software, including the TWAIN driver and use it, if you prefer it.

Attention: There is a less-popular competitor to TWAIN called Image and Scanner Interface Specification (ISIS) that you may occasionally encounter support for in some applications. It allows not only image scanning but also image handling and processing. ISIS driver for most scanners are available at www.scannerdrivers.com.

Do not confuse the TWAIN driver with the image-editing software that comes with most scanners. The editing software may be TWAIN compliant (and probably is), but the ability to acquire images directly from the scanner is a function of the TWAIN driver, not of the image-editing program.

2.7.2 Windows Image Acquisition (WIA)

As we mentioned earlier, Windows XP and Me support an Application Programming Interface (API) called Windows Image Acquisition. It works with

both scanners and digital cameras and allows users to connect supported devices without TWAIN, ISIS or any other special software.

When you scan with a scanner's own software through TWAIN the interface and options vary from scanner to scanner and even from version to version of the software. Therefore, if you work with several different scanners, there is a learning curve involved. With WIA, however, every scanner can be accessed through the same Scanner and Camera Wizard interface in Windows, so you know exactly what you are doing at all times. Figure shows the Scanner and Camera Wizard from Windows XP at work with a HP multifunction device that includes scanning capability, for example.

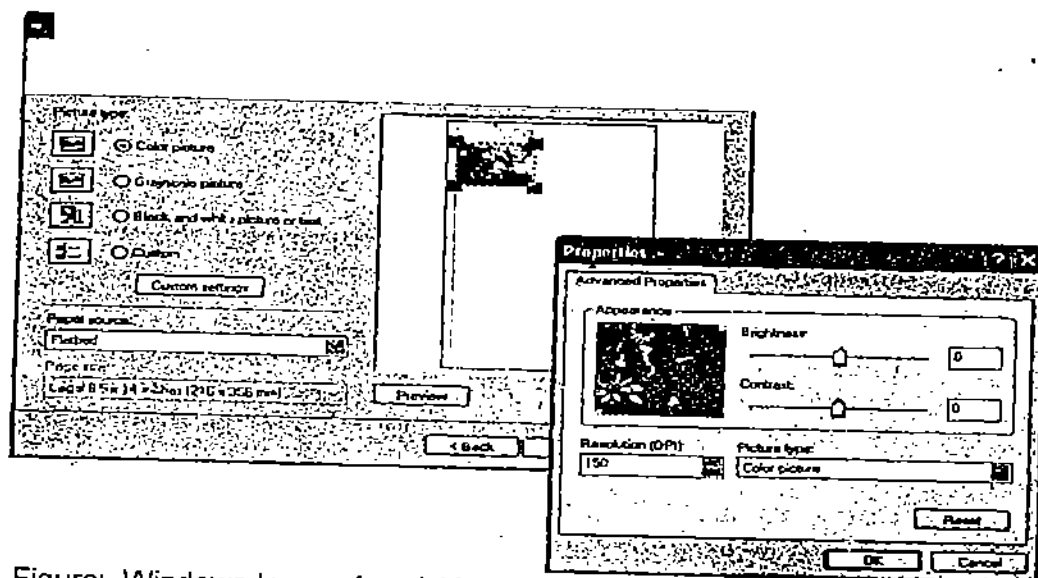


Figure: Windows Image Acquisition offers a consistent Scanner and Camera Wizard interface for all scanners.

Hint: WIA was new in Windows Me and was still getting the bugs worked out; you will find that it works much more smoothly and with a much wider variety of hardware under Windows ME and XP.

2.7.3 Optimization of Size and Quality of Scan Image

As you are scanning, you will have a choice of the resolution at which the scan occurs (for example, in figure it is set to 150 dpi) and of the format in which the scanned file will be saved. These are both important decisions. For the best scans possible, you may have to change the resolution, the file type, and perhaps some other settings in the scanning application as well.

2.7.4 Selection of Appropriate Image Resolution

If you have extra money for a scanner with a very high maximum resolution (dpi), you may be tempted to, always use the highest available resolution setting. After all, more is better right.

Well, sort of. Higher resolution may result in a better quality scan, but it also results in a huge file that takes up many megabytes on your hard disk, is impossible to distribute to others via e-mail, and takes forever to download when

placed on a Web site. And, when viewed on a computer monitor, the higher resolution is completely wasted because the average monitor displays at 72 dpi.

You might want a high-resolution scan, if you are scanning something that **will** eventually be placed in a professional-quality publication and printed. In a **case** like that, you will want to match the resolution of the picture to the maximum resolution of the printer in which you will be printing. For example, if you are going to be printing to a commercial-grade full-editor printer that prints at 1200 dpi, then scanning at 1200 dpi might produce a noticeably better image than, say, 300 dpi.

However, if you are scanning something that will eventually be placed on the Web, you want to keep resolution as low as possible without noticeably degrading the image. People visiting the Web page will (or should) thank you for keeping the image size small so that they do not have to go and get a cup of coffee while the page is loading. For an image that will be viewed only on screen, 75 dpi is adequate. Scan it at 150 dpi in safe side.

2.7.5 Selection of Appropriate File

The file type in which you save the image also makes a big difference in the file size. Avoid using BMP format because it is bloated and not particularly stellar in any way. TIF is great for high-resolution images where printed quality is important, but it results in large file sizes. For Web use, go with the JPG format, which is quite compact and results in decent quality images good enough for on screen viewing.

Crop Images: If you do not need the whole image, do not scan the whole image. The Scanner and Camera Wizard (WIA) and most TWAIN-compatible scanner interfaces allow you to preview the image before doing the "official" scan, and after the preview, you can adjust crop marks so that only part of the scanner glass is scanned. If that is not possible, then use the Crop feature in your image editing software to cut it down to size. Cropping only what you need from an image not only makes the file size smaller but also makes for a better-looking picture in many cases.

2.7.6 Selection of Appropriate Image Mode

Most scanner interfaces enable you to choose between at least three settings:

Color: Grayscale and Black and White (line art). This is an important decision with implications for both image size and image quality and color. This is the default usually. It scans in full color and results in the largest file size.

Grayscale: this scans in 256 or more levels (shades) of gray. This results in a smaller file size than color is useful when you know you are going to output to black-and-white media (such as a black-and-white printed newsletter or newspaper).

Black and White this scans is only one color: black. It results in the smallest image size, but all shades and colors are converted to either black or white, resulting in something like a fax. This is useful for line art but very poor for photos.

2.7.7 Optical Character Recognition

Many scanners also include Optical Character Recognition (OCR) software, which enables you to convert scanned text into computer-based text. This way you can scan a document and import it directly into a word processor (such as Microsoft Word) as editable text, rather than as a graphic.

Most current tilt OCR software does a good job of translating printed characters to digital characters, but you will still need to clean up any misinterpretations. The cleaner the scan, the better the job the OCR software does, so start with a clean original and make sure you take a good, high contrast scan. (Also, be sure the original document is centered and level on the glass bed and has no creases or blemishes.)

Attention: Sometimes advanced OCR software referred to as page recognition software also captures fonts and page layout from the source document.

Hint: OCR software works best when capturing black-and-white images. All the extra information in a color image only serves to cut if use the software.

2.8 Scanner Maintenance

Scanners do not require a lot of maintenance. (Unlike printers, scanners do not use up any consumables, such as paper or ink.)

The main thing is to keep your scanner clean and as dust-free as possible. Use a soft cloth and glass cleaner to clean the glass bed of scanner. Take particular care to keep the glass free from smudges and scratches, both of which can affect the quality of your scans. You may also need to disassemble it to clean the under side of the glass, which can also get dusty.

As to mechanical maintenance, most consumer level scanners do not have too many moving parts to which you have easy access. Depending on your scanner, even the light source (typically a lamp or small light bulb) may or may not be serviceable. Many manufacturers now use sealed lamps that cannot be replaced by the user.

If you have a sheet-fed scanner (or an all-in-one unit that includes a sheet-fed scanner function), make sure the paper path is clean and free from jams and paper bits. Do not inadvertently jam the paper path by force-feeding paper that is too thick to fit.

The truth is given the low price of today's scanners, if it breaks you will probably be tempted to just replace it.

2.8.1 Common Problems and Solutions

Before the dawn of USB, most scanner problems came from trying to share a parallel port with the printer. If you have a USB scanner, you automatically have an easier installation and fewer subsequent problems. Beyond the parallelism, other common scanner problems include the inability to scan and poor quality uses.

Check the Connections: As any peripheral device connected to your system, there are some basic things; you need to check if you are having trouble with device.

First, make sure the scanner is plugged into a live power supply and is turned on. You should see a light below the glass bed when the scanner is turned on and operational.

Second, make sure the scanner is properly connected to the correct port on your PC. Check both ends of the cable, just in case may be.

Finally, many scanners ship with their CCD and stabilizer arm locked down for transport. The scanner would not do much scanning if the stabilizer arm cannot move, so check the manufacturer's instruction manual to make sure you have unlocked the CCD properly. A horrible grinding or clicking noise, the first time you try to use the scanner is a good indicator of a locked CCD.

Test: When you are trying to track down a problem with your scanner, it helps to follow a systematic troubleshooting approach. For that reason, the first thing you should do (after you have the scanner connected and powered up) is performing a comprehensive test of your scanner to determine the nature of the problem..

TIP: Many scanners have a "self-test" mode that you can run to check out the unit's operation. In addition, some scanners come with software-based diagnostic utilities.

Here is what you need to do:

- Turn on your scanner, and launch the scanner's image acquisition utility or software or in.
- Windows XP or Me (with a compatible scanner) run the Scanner and Camera Wizard from the Accessories menu.
- Place a black-and-white photo facedown on the glass bed. (If you are using a sheet-fed scanner, insert a black-and-white document into the sheet fed and make sure it is facing the correct direction).
- Acquire the image with the scanning software.
- If the scanned image is blank (either all black or all white), make sure the original photo or document was either placed facedown or inserted proper side up into the sheet feeder. (You could be scanning the back of the photo)
- If the acquired image is blurred or distorted, make sure the original photo or document was set firmly in place and that the cover of the scanner was closed. You may also need to select a particular type of scan in the scanning software. (You will probably have choices such as B/W Photo, Color Photo, Line Art, and so on). Choose the most appropriate image type.
- Once, you get a good preview image, save the scan to disk, experiment with scanning the document at different resolutions and image settings.
- Now it is time to check the TWAIN driver (if you are even using TWAIN; you might not be you have Windows Me or XP). Close the software of scanner and open another graphics program, one that you know is TWAIN compliant. Select File Open or File Acquire, and select your scanner from the Select Source option. Choose to acquire the image, and

see if the graphics program actually acquires the image from your scanner:

- If you cannot acquire an image via the graphics program, you should check several things. First, make sure the program is TWAIN compliant; if not, use another program. Second, if your scanner does not appear on the Acquire Source menu, you may not have installed the TWAIN driver; try reinstalling the driver and testing again.

Hint: If, no matter how hard you try, you do acquire scanned images from within your graphics program, you can always default to save the images to disk using your scanner's acquisition software and then opening or importing those files into your graphics-editing program as normal.

Check the Port and the Configuration: If you have connected your scanner via a parallel port, make sure your port is configured as an Enhanced Parallel Port (EPP) or Extended Capabilities Port (ECP). Many scanners require EPP / ECP operation, which utilize high speed, bidirectional data transfer to send data back and forth to your computer.

If Windows is not recognizing your new scanner, your best course of action is to rerun the play that is disconnecting the scanner and try reinstalling it. If Windows does not recognize the scanner when you reconnect it, try uninstalling the scanner, turning off your computer, and then reconnecting and rebooting. Windows should recognize the new device when it starts up again.

If Windows still does not recognize your scanner, run the Scanner and Camera Wizard (in Windows XP) or the Add New Hardware Wizard (in any version of Windows). Either of these options should let you install the scanner driver manually.

With some scanners, an even better approach is to install the unit from the installation CD of scanner. Some scanners include a robust installation program that will do everything from installing the drivers to configuring your system to installing the scanner's graphics editing software.

Hint: When in doubt, read the scanner manual for the recommended installation method or visit the Web site of manufacturer for the latest drivers, installation instructions and advice.

If your system recognizes your scanner after installation but does not recognize it on subsequent use, you may have a configuration that requires your scanner to be powered up before Windows starts. That is, Windows will not recognize your scanner if you turn it on after Windows is up and running. Power down your computer, turn on your scanner, and then power up your computer again. When Windows loads, it should recognize the already running scanner.

A related problem occurs when your scanner is set too far away from your computer. You should never use a cable longer than the one supplied with your scanner; long cable runs will not only degrade the image quality but will also cause your computer not to recognize the signals coming from the scanner. Under any circumstances, do not use an extension cable to lengthen the cable from scanner to PC.

With a USB scanner, if you are having problems connecting through a USB hub, try connecting directly to the USB port on the PC. It should not make a

difference with a powered USB hub, but with an unpowered one (such as a USB extension port on a USB keyboard), it might.

Allocate Necessary Resources: Scanners can be resource hogs. Not only do scanned images take up a lot of disk space, scanners also use a lot of memory and processing power when they are acquiring images.

If you find your system slowing down or hanging when you try to make a scan, consider the following options:

- Close down any unnecessary applications while you are scanning.
- Add more Random Access Memory (RAM) to your system.
- Make sure you have plenty of hard disk space available because most scanners take full advantage of your swap file of system; you can also increase the size of your swap file of Windows for better handling large scanned images.

If you still run into performance problems, try scanning at a lower resolution. (Smaller pictures use fewer resources).

Troubleshoot Parallel Port Sharing Problems: When a scanner and a printer share a parallel port, there are some inherent issues. Thankfully, this is a less common situation than it was in the past because most scanners are USB nowadays. However, we will tell you about it anyway because as you know, there is always the occasional client who is still working with old equipment.

If you try to use both devices at once, you will either slow down both devices or freeze up one or both of them. You should avoid scanning while printing (and vice versa) or just install a switch box that only lets you feed one device at a time to your PC's parallel port.

You should also consider the order in which you power up each device. You should always turn the printer on first, then scanner, and then your computer. Many manufacturers provided that scanner be switched on and connected for the printer to work.

Finally, your printer might not be configured to share your PC's parallel port. Open the printer configuration utility and look for a setting like use port exclusively. Then turn that setting off. Also, deactivate any Use Bidirectional Printing setting. Changing these settings will allow other devices (such as your scanner) to share the parallel port with the printer. You might also try some different parallel port mode settings in Basic Input/Output System (BIOS) Setup.

If you happen to have a spare parallel port Input/Output (I/O) board, you can install that in your system to give yourself LPT2. That way the devices do not have to share. We would not buy a parallel port I/O board for this purpose, however, because for \$50 you can buy a cheap USB scanner and avoid the issue entirely.

Some Fix Problems of Scan Quality: If a scan is not like what you expected, if it is too light or too dark, crooked, or whatsoever you do not have to rescan the image. You can fix many simple problems via the scanner's graphics editing software or in any graphics-editing program.

If your scan is too dark or too light, simply adjust the brightness and contrast in the graphics editor. If the scan is crooked (a common problem), use your graphics editor to rotate the image a few degrees. If the scanned image is off-

center, crop the unnecessary portion of the image in the graphics editor. In other words, you can correct many scanning mistakes.

If, on the other hand, you see vertical streaks in your scanned image, your problem is with the scan and it must be fixed there. These streaks are caused when your scanner scans a dust or dirt particle along with the original document. To eliminate this type of image streaking, make sure the glass bed is clean (on a flatbed scanner) or that the paper path is clear (on a sheet-fed model).

If you see circular diffraction patterns in your scans, you probably have a moisture problem with the original item, particularly if you are scanning a slide or film negative. These Newton rings are created when the scanner's light hits the moisture on the original item. You can eliminate the rings by eliminating the moisture or by using a film or slide holder to lift the original off the scanner's glass surface.

Another common problem is the presence of a pattern over halftone or patterned areas of the image. This is caused when the dots used to create the original printed picture are smaller than the pixel size of your scanner. You can fix this problem by increasing your scanner's resolution (thus making the pixels smaller) or by shrinking the scanned image, which makes the pixels more closely match the original's dot size of image.

Hint: If a scanned image does not look right on your monitor, it may not be a problem with your scanner. If you have your monitor set for a low bit-rate display 256 color, for example, you might think that you are viewing a bad scan, when, in fact, your monitor is causing the apparent problem. That is because your scans are probably made with 24-bit color and thus do not reproduce well on your monitor if it is set for a lower bit rate. To fix this problem, open Windows Display Properties dialog box and increase the bit rate of your monitor to 24 bits or more.

2.9 More Tips on Troubleshooting

The following is a summary of troubleshooting tips for your scanner:

- If your scanner is connected via the parallel port and is not communicating with your computer (or if your scanner and printer are interfering with each other), check the order in which everything is powered up (scanner first, printer second, PC third) or consider installing a parallel port switch or a second parallel port card (LPT2).
- If your stabilizer arm of scanner is not moving, unlock it. (The lock is typically located on the bottom of a flatbed scanner.)
- If you cannot acquire images from your scanner, reinstall the scanner's TWAIN driver.
- If your scans are slow and you are connecting via the parallel port make, sure the port is configured as either an EPP or ECP.
- If your scans are slow (no matter how you are connected), consider increasing the size of your PC's swap file or adding more RAM to your system.
- If your scans are coming out blank, make sure you are placing the original facedown on a flatbed scanner or in the appropriate direction in a sheet-fed scanner. If the scanned image is blurry or distorted, make sure the

original document is placed firmly in or on the scanner and that, the flatbed scanner's cover is closed.

- If the quality of your scans is poor, make sure your scanner or scanner software is set to the appropriate setting for the type of image you are trying to scan.
- If you find minor errors in your scan, it has crooked, off-center, or too dark or too light use your scanner software or graphics-editing software to fix the errors in the scanned image.

In general, here is what you want to remember: if you can, connect your scanner via USB instead of the parallel port. Make sure the item to be scanned is in a good condition and centered on the glass bed or in the sheet feeder. (While you are at it, make sure the glass surface is clean and the sheet feeder is clear.) Set your scanner or scanner software to the appropriate setting for the original image type. Then use your scanner software or image-editing software to edit any time mistakes that pop up in the scan.

Follow that general advice, and you will minimize the opportunity for scanner related problems.

2.10 Summary

In this unit, we have discussed about an important input devices like scanner. We have also described installation and configuration of this input device. We have introduced different components of scanner. Finally, we have explained about trouble shootings and maintenance of scanner.

2.11 Model Question:

Write a short note on the following.

1. What are roles of scanner? Can you explain its components?

2. Give the comparative picture of all type of scanner used in computer field.

3. Write the installation and maintenance scanner. Describe image quality under Windows environment.

UNIT-3: INSTALLATION AND MAINTENANCE OF CD AND DVD DRIVES

Structure:

- 3.0 Objectives**
- 3.1 Introduction**
- 3.2 Installation of CD and DVD Drives**
- 3.3 Function of CD Drives**
- 3.4 Types CD and DVD Drives**
 - 3.4.1 CD ROM (Compact Disk Read Only Memory)**
 - 3.4.1.1 CD-Recordable (CD-R)**
 - 3.4.1.2 CD-Rewriteable (CD-RW)**
 - 3.4.2 Digital Versatile Disc (DVD)**
 - 3.4.2.1 Recordable or Rewriteable DVDs**
 - 3.4.2.2 DVD-R**
 - 3.4.2.3 DVD-RW**
 - 3.4.2.4 DVD-RAM**
 - 3.4.2.5 DVD+RW**
 - 3.4.3 CD Drive Standards**
 - 3.4.3.1 Red Book**
 - 3.4.3.2 Yellow Book**
 - 3.4.3.3 HIGH SIERRA AND (ISO 9660)**
 - 3.4.3.4 CD-ROM/XA**
 - 3.4.3.5 MULTISESSION CAPABILITY**
- 3.5 CD / DVD Drive Connection**
 - 3.5.1 EIDE**
 - 3.5.2 SATA**
 - 3.5.3 SCSI**
 - 3.5.4 USB**
 - 3.5.5 IEEE 1394 (FireWire)**
- 3.6 Characteristics**
 - 3.6.1 DATA TRANSFER RATE**
 - 3.6.2 ACCESS TIME**
 - 3.6.3 CACHE/BUFFER**

3.7 Steps of Installing CD or DVD Drive

3.7.1 Preparing a Drive Bay

3.7.2 Setting Jumpers

3.7.3 Installing the Drive in the Bay

3.7.4 Connecting the Cables

3.7.5 BIOS Configuration

3.7.6 Drivers

3.8 Trouble shooting of Drives

3.8.1 Recording CDs and DVDs

3.8.2 Avoiding Buffer under runs

3.8.3 Recording CD and Data DVDs

3.8.4 Producing and Recording DVDs

3.8.5 Selecting Writeable Media

3.8.6 Handling CDs and DVDs

3.9 Summary

3.10 Model Question

3.0 Objectives

At end of this unit, student will be able to:

- Understand about the basic function of CD, DVD and their drive
- Explain the various kinds of CD, DVD that are available in the computer.
- Distinguish in between CD and DVD with respect to their feature.
- Know that how to perform read and writing operation by the CD and DVD.
- Concept various types of interfacing devices used in connection.
- Analyse statistics of capacity of CD and DVD.
- Generate awareness about different standard of CD drive available in the computer.
- Identify various trouble shootings keyboard and mice.

3.1 Introduction

Almost all software comes on Compact Disc Read-Only Memory (CD-ROM) these days, so having some type of CD drive is nearly a necessity. However, which kind is best? A writeable CD drive, either CD Recordable (CD-R) or CD Rewriteable (CD-RW). Similarly, a Digital Versatile Disc (DVD) drive, either a writeable DVD or some sort of combination. And, if you want a writeable DVD, which of the many competing standards will you may select. This unit attempts to provide some answers to those questions by explaining CD / DVD technology and how to install, use and maintain optical drives and discs.

3.2 Installation of CD and DVD Drives

No matter what kind of CD drive you are installing (including DVD); the physical installation of the drive is physically the same. DVD drives require a little bit of extra cabling if you are also installing an MPEG decoder card (required to watch DVD movies on screen), but other than that, they are similar.

To install a drive, follow these steps:

1. If the drive requires a separate adapter card, install it in the PC. Most CD drives run off the motherboards existing Integrated Device Electronics (IDE) interface, so this is probably not necessary.
2. If you are installing, a DVD drive and you have an MPEG decoder card to install, attached it in the PC. These are Peripheral Component Interconnect (PCI) cards.
3. Set any jumpers on the drive as needed fits an IDE drive you are installing set the appropriate Master, Slave or Cable Select (CS) settings. If SCSI, set the SCSI ID number and or termination as needed.
4. Install the drive into an open bay.

5. Attach the power cable to the drive.
6. Connect the interface cable to the drive interface (on the motherboard or adapter card) and to the drive.
7. Attach the sound cable from the CD drive to the sound card if you want to be able to dial audio CDs.
8. If you are installing an MPEG decoder card as part of a DVD installation, attach a loop back cable from your computer's video board to the DVD decoder Video In port, and then connect your monitor cable to the DVD decoder Video Out connector.

3.3 Function of CD Drive

CD and DVD drives fall into the optical category of drives. The installation of Hard Disk Drive and its terminology is same as hard disks, floppy disks and Zip disks are all magnetic in nature. In other words, they store data in transitions between positive and negative magnetic charges on the disk surface. Optical discs and drives, on the other hand, store data in patterns of transitions between areas of greater and lesser reflectivity on the shiny surface of a disc.

When you insert a CD or DVD into a drive, a sensor detects the disc, and the spindle motor begins spinning it. A laser shines light on the surface of the disc, and the shiny surface reflects light back to a photo detector sensor. The amount of reflected light is measured and that data is sent to the drive controller.

Data on an optical disc is stored in patterns of pits (or the appearance of pits) on the disc's surface. A pit is a tiny indentation in the surface that causes the light to be reflected back less strongly than in other areas. The non-pitted areas are known as land.

It is to think that a pit stands for a binary digit, but that is not how it works. Binary digits are indicated by the patterns of transitions between pit and land areas, much like on a magnetic disk. When the sensor detects a change for light bounced back, which could be a transition from either land to pit or from pit to land; it sends an electrical pulse to the controller that indicates a 1 bit. When enough time passes that, a transition could have occurred but did not, it interpret that as a 0 bit.

Mass-manufactured CDs and DVDs have actual pits in the surface, covered by a protective clear coating. Writeable CDs and DVDs work a bit differently, as you will learn later in the unit. They use a writing laser to create as of lesser reflectivity on the disc's surface that simulates the same loss of reflectivity that a pit would create. Therefore, the photo detector sees those areas as pits.

Now that you know how the drives work, you will look at the differences between the many different optical drive and disc standards available today.

3.4 Types of CD and DVD Drives

Today there are many different types of optical drives available, but the basic level is CD-ROM. Regular CD-ROM drives can read from standard data or audio CDs, but they cannot read DVDs and cannot write to any disc.

When we are referring to optical discs, we use disc; when referring to magnetic disks, we use disk. However, not everyone uses this convention.

All of these other types can also read standard CDs in addition to their own special capabilities:

- CD-R drives can write to write-once CD-Rs.
- CD-RW drives can write either to CD-Rs or to multi-rewriteable CD-RW,
- DVD-ROM drives, also called DVD players, can read from data DVDs and can play DVD movies (with the help of an appropriate MPEG decoder).
- DVD-recorder drives can do everything a DVD-ROM drive can do and they can write to DVD-R (write-once) or DVD-RW (multi-rewriteable) discs.
- Combination drives have the capabilities of more than one of the previous types. For example, a combination drive might include CD-RW and DVD-ROM capabilities.

3.4.1 CD-ROM (Compact Disc Read Only Memory)

You will see some kind of CD drive in virtually every computer sold today, and in lower end or older systems, it may be a regular CD-ROM drive (that is, one that lacks any of the special qualities such as writing CDs, writing DVDs or reading DVDs).

Regular CD-ROM drives read data CDs, just as you read the contents of any older disk on a PC. They also play audio CDs, provided you have a CD player program installed on your PC. They are the baseline of functionality; all DVD and CD-RW and DVD-RW drives perform these functions, so you do not need a regular CD-ROM drive if you have one of the other types.

So generically speaking how does a CD store data. Well, the surface of a regular CD is arranged as a single spiral that begins at the inside of the disc and travels outward toward the edge of the disc. The commercially produced ones are made of injection-molded plastic formed from a stamper disc, coated with an aluminum film to make them reflective and then lacquered to protect the surface.

From 527 megabytes (MB) to 742MB will fit on a disc, depending on the number of sectors on the disc and the format used on it. Standard discs use the same size spiral as an audio CD (60 minutes) with 270,000 sectors and up to 99 tracks. The typical disc used in a computer-based CD-ROM drive stores 630MB using 333,000 sectors.

CDs are physically organized a little differently than hard disks. Whereas hard disks have individual tracks in concentric circles, CDs define a track by the length of a specific file; one file equals one track. The tracks on CDs are having sequentially

on a continuous spiral. This spiral is a staggering 3 miles long and can contain more than 2 billion pits.

Commercially produced CDs are made up of a polycarbonate (plastic) wafer, about 12 millimeters (mm) thick. This after is coated with metallic film, which is then covered, for protection, by a plastic polycarbonate coating. Most CDs are single-sided; the player reads data normally in one side. (In the early days of CD technology, there were a few double-sided discs, but you will seldom run into one today.)

Regular CDs store data by arrangements of pits on the CD (about 0.12 microns deep and 0.6 microns wide). The pits produce areas of less reflectivity, so when a laser hits the disc, the light does not bounce back as strongly. Just like with 1s and 0s in binary computer code, this "pit or no pit" system stores binary data on the CD, as we said earlier.

3.4.1.1 CD-Recordable (CD-R)

CD-R and CD-RW are two technologies for recording your own CDs. CD-R, the older technology writes only once to a blank, writeable CD. After you write to the CD-R, You cannot write to it again (except in the case of multisession disc, which I will get into later). It is great for data archiving or creating music CDs. Recall from the preceding discussion that regular CD-ROMs are composed of metallic film on a plastic disc, and the disc is pitted to store the data by reflecting or not reflecting light from the pit or non-pitted areas. On the other hand, commercially available CD-Rs are coated with metal in a process, similar to electroplating and overlaid with photosensitive organic dye. When the laser "burns" data into the CD-R media, it heats the metal and the dye together, causing the area to diffuse light in the same way that a pit would on a mass-produced CD so that CD-ROM drives are fooled into thinking that a pit exists. There is no pit, though just a spot of less reflectivity.

CD-R has its drawbacks. If the recording process falters, you have just wasted a blank CD (they cost about Rs. 10 a piece these days, or less if you buy in bulk). CD-R is also not great for backing up data that change frequently because you have to use a new CD every time something changes to make sure you have current backup.

3.4.1.2 CD-Rewritable (CD-RW)

A CD-RW drive can write to a blank, rewriteable CD multiple times, just like with a hard or floppy disk. You can erase files stored on a CD-RW and make changes to them, so CD-RW is an ideal format for backing up constantly changing data, such as data files from an accounting program you use daily. You can write to a typical CD-RW more than 1000 times before it wears out. Although, CD-RW blanks are more expensive, they last longer because you do not have to throw one away when the data on it becomes outdated. Most CD-R W drives sold today are dual-purpose drives that can do CD-R as well as CD-RW, depending on the type of blank you insert in them.

On a CD-R, the dye changes into a non-reflective state because of mixing with the underlying metal. That change is permanent, which is why you cannot alter a CD-R's content. In contrast, CD-RWs also contain material that changes reflectivity when exposed to a laser, but not permanently. The CD-RW contains a silver, indium, antimony, tellurium alloy that is naturally reflective. When the CD-RW drive writes, the laser uses a high-power setting to heat the alloy to about 500 to 700 degrees Celsius ($^{\circ}\text{C}$), causing it to liquefy. After being liquefied, a spot loses its reflectivity when it hardens again, making an artificial pit. By reheating the area to a lower temperature (about 200°C , a setting called Erase), the laser makes the material revert to its reflective state. In this way, it is possible to write to CD-RWs many hundreds of times.

Before you can use a CD-RW, you have to format it much like you format a hard disk, but with a special utility, that comes with the CD-RW software. Early CD-RW drives took a long time to format discs (more than 30 minutes in some cases), but modern drives can format a CD-RW in about 30 seconds. Formatting lays down a grooved spiral on the disc, ensuring accurate alignment of the burned-in data.

Even though CD-ROM, CD-R, and CD-RW discs look similar, they store data differently as we have just explained; that is why some older CD players cannot read CD-R and CD-RW discs. All the newer ones can, though.

3.4.2 Digital Versatile Disc (DVD)

Like CD, DVD is known by its initials. The DVD format squeezes more data onto a single disc than is possible with a regular CD. CD-ROMs can fit about 630MB onto a disc whereas DVD can fit as much as 17 gigabytes (GB). DVDs are used to distribute movies, and increasingly game manufacturers are releasing games and business applications on a single DVD that would otherwise have required multiple CDs.

A DVD is somewhat like regular CD technologically. DVDs can store more information than regular CDs because they have a smaller pit length, less space between pits, a slightly larger data area (less wasted space around the outside and inside) and a few other minor design improvements. The primary difference is that the pits are much smaller and much closer together.

There are four kinds of DVDs: single and double layered and single and double sided. You might hear certain DVDs referred to by a number such as DVD-9 or DVD-10. That is a reference of the standard they support. The following are some of the popular standards:

- DVD-5: single-sided, single layered 4.7GB capacity.
- DVD-9: single-sided, double-layered, 8.5GB capacity.
- DVD-10: double-sided, single-layered, 9.4GB capacity.
- DVD-18: double-sided, double-layered, 17GB capacity.

The second data layer is written to a separate substrate below the first layer, which is semi reflective to allow the laser to penetrate the top layer to read the

substrate below. This enables the disc to hold more than twice the data on the same surface area. On a double layer disc, the different layers have different colors. Therefore, a double-layer disc may appear to have a gold color to it whereas a single layer one looks more purely silver.

Attention: Another term you might occasionally hear associated with DVDs is RSDL, which stands for Reverse Spiral Dual Layer. This refers to the ability for the player to automatically go to next layer at the end of the first one. This is useful when playing a movie because it does not interrupt the playback for example.

A DVD drive can also function as a regular CD-ROM drive. Therefore, you do not need to have both in your PC; a DVD drive can replace your existing CD-ROM drive. In fact, DVD-ROM technology is on its way to replacing CD-ROM technology on the computer as a way to distribute software and high capacity data.

3.4.2.1 Recordable or Rewriteable DVDs

If a writeable CD is good, then a writeable DVD must be even better. Not only can it hold hours of video footage, but it can also hold huge quantities of data, so a DVD seems like the perfect medium for large backups and data storage.

The computer industry has been enthusiastically pursuing the development of writeable DVD technology in recent years. Unfortunately, all this enthusiasm has produced many different standards. The situation you have is similar to the early days of VCRs. Except that, recordable DVDs are available in four basic physical formats:

- DVD-RAM
- DVD-R
- DVD-R W
- DVD+RW

There are two types of DVD discs in terms of what is stored on them: DVD-video and DVD-ROM. DVD-video is what you buy at your local video store. These discs are optimized for the storage and playback of digital movies through a player connected to your television. These discs are analogous to audio CDs that you play in your stereo. DVD-ROM is the computer counterpart and, as such, is more similar to a CD that you use at your computer every day. The DVD-ROM drive in your computer will play both type of disc. However to play DVD-video on a computer, you need either dedicated software or a special adapter card called an MPEG2 decoder to decode the movies.

Some DVD drives come with the needed decoder card; others make you buy it separately. Using such a decoder board is known as hardware decoding. The alternative is to use a software driver that does the decoding. This is software decoding. Hardware decoding results in better quality video playback, but on a powerful PC with lots of Random Access Memory (RAM) there is not much difference. The need for hardware decoding was mostly an issue in the past when PCs were not powerful enough to handle file decoding on their own through software; however, some video files continue to prefer them for the highest quality movie playback. Nearly all mainstream desktop PCs today go with software decoding.

None of these able formats are fully compatible with one another, and there are even compatibility problems with existing drives and players. Using the old Beta versus VHS VCR analogy again, you could not play a Beta tape in a VHS player, and you could not play a VHS tape in a Beta player.

When you look at the specifications for current DVD recorders, you see a mixed bag of compatibility, so you will have to research different recorders before you make your purchase. If you want to Use the DVD recorder for your own system backups and data storage, any of the following formats will probably suit your needs. However, if you intend to share or distribute your DVDs, you should select a recordable DVD format that supports the needs of your intended audience:

3.4.2.2 DVD-R: First available in the fall of 1997, this popular format comes in two versions: DVD-R for general use and DVD-R (A) for authoring. The general version uses a laser that is more compatible for the future ability to write DVD-RAM. DVD-R (A) is intended for professional development and is not writeable in DVD-R drives. Both types, however, are readable in most DVD players and drives. Similar to CD-R technology, this format records sequentially to the disc and allows only one recording. The capacity of a first generation disc was originally 3.9GB and later increased to 4.7GB.

3.4.2.3 DVD-RW: Introduced in Japan in December 1999, this rewriteable format was created by Pioneer. It can be rewritten about 1000 times. DVD-RW is based on the DVD-R format and can record from one to six hours of video depending on quality. DVD-RW discs have 4.7GB capacity. Though it has not usually required that you use a caddy or cartridge for the DVD-RWs, some drives may require one. DVD-RWs are playable in most DVD drives and players. However, when you insert a DVD-R W into some DVD-video players, the player may assume it is trying to read a dual-layer disc.

3.4.2.4 DVD-RAM : Introduced in the summer of 1998, this rewriteable format has many manufacturers behind it and is well suited for storing computer data. However, DVD-RAMs aren't-compatible with most existing drive and players because of format differences. This format is also considerably more sensitive to handling than other formats. A double-sided DVD-RAM must be enclosed in a cartridge, similar to the caddy used with early CD-ROM drives. For single-sided DVD-RAMs, the cartridge is optional but recommended. Its initial capacity was 2.6 GB per side, for a total of 5.2GB. Second-generation DVD-RAMs has a capacity of 4.7GB per side and are more suitable for editing and accessing movies and music. Second-generation drives are backward compatible with 2.6GB DVD-RAMs and can read other DVD and CD formats. DVD-RAM is a rewriteable format that can handle up to 100,000 rewrites on a single disc, which is more than 100 times that of DVD-RW.

The main problem with the DVD-RAM format is with video. Most home DVD players will not play movies on DVD-RAMs. However, some DVD-RAM recorders will also record in DVD-R format as well (a write-once format) for compatibility.

3.4.2.5 DVD+RW: DVD+RW are are-writeable format developed as a joint effort by the DVD+R W Alliance. Members include Hewlett-Packard, Mitsubishi Electric,

Philips, Ricoh, Sony and Yamaha. It offers fast access and rewrite ability like DVD-RAM but without that format's incompatibility issues. DVD+R W is not an officially sanctioned standard at this point (so buyer beware), but it does seem to offer the best of both worlds of DVD-R and DVD-RAM.

Hint: When buying a DVD player for your television system with homemade DVDs in mind, look for one with the DVD-Multi certification. These players should play discs in DVD-R/RW and DVD-RAM format. When, burning DVDs for use in a DVD movie player, use the write-once discs (DVD-Rs) rather than the rewriteable ones for the best chance at compatibility.

3.4.3 CD Drive Standards

CD drive standards used to be important in the early days of the technology because there were all sorts of competing technologies that were not compatible with one another (just like now with the writeable DVD standards). However, nowadays the standards for regular CD-ROM are well established to the point where looking for compliance with particular standards is no longer significant. Nevertheless, you might find it interesting to read a brief overview of the CD drive standards that developed over the past decade or so. If not, free to skip this section.

A book color system was used to designate CD primary standards: red, yellow, orange, green and white. The original specification was published in a book with a red cover, so that is how the naming convention originated. The next book they published with a yellow cover, and so on. Each color stood for the standards for a particular usage; therefore, they do not overlap or contradict one another.

COLOR	TOPIC
Red	Audio
Yellow	Data storage
White	Video CD (movies)
Orange	Writeable CD (CD-R and CD-RW)
Green	Combination of Red and Yellow

Table: CD STANDARDS BY BOOK COLOR

It is important to note that these colored book standards are not frozen in time. The Yellow Book standard is not one single document, for example; it has been updated many times over the years to accommodate new technological improvements. Rather than, each book color represents a certain functionality of CD drives.

3.4.3.1 Red Book

The Red Book was the original specification defining how digital audio information is stored and indexed. It was developed for music CDs, which were the first CDs commercially available, and almost all music CDs conform to this standard today. All

CD drives available for computers today support audio CD playback by conformance to this standard.

3.4.3.2 Yellow Book

The Yellow Book also named ISO 10149 (ISO stands for the International Organization for Standardization), extends the Red Book audio specifications and deals specifically with the more interactive requirements of CD-ROM: random access capability and multimedia. The specifications concentrate on storing and indexing data and error correction.

This standard supports several few formats for the different computer platforms that use CD-ROMs. These include Native Macintosh hard disk format, Digital Equipment Corporation's (DEC's) VMS, Apple's Hierarchical File System (AHFS), and ISO 9660 (which you will learn about next). As we pointed out earlier, though, these standards are a point of interest mostly for their historical value today because Yellow Book technology has become standardized.

3.4.3.3 HIGH SIERRA (ISO 9660)

In 1985, CD-ROM companies formed an alliance to produce a standard CD file structure independent of the operating system. The result is commonly known as the High Sierra standard. After some modifications, the standard was formally accepted by the ISO and was named ISO 9660. It sets forth a standard file system with a hierarchical directory structure of eight levels. This Compact Disc File

System (CDFS) is similar to the File Allocation Table (FAT) architecture familiar to DOS users. It is the standard format for all data CDs produced today, when you look at a CD drive's properties in an operating system (Windows, for example), you will see the file system reported as CDFS; now you know what that means.

3.4.3.4 CD-ROM/XA

The CD-ROM/XA standard was another extension to the Yellow Book. The XA stands for Extended Architecture, and this standard does exactly that. By defining the way in which different data types may be interleaved (that is, woven together) on a CD, the XA standard really makes multimedia CDs possible. It also allows for multisession recording.

A common application requiring the interleaving of sound and motion can be demonstrated through a multimedia presentation. Suppose you are giving a presentation with pictures of a person speaking and have sound to follow along. In order for the movement of the speakers lips to match the spoken words, the sound and the pictures must be synchronized. To accommodate this, tracks on a CD-ROM/XA can contain interleaved video, audio, and computer data.

Briefly, here is how it works: a standard CD track contains only Mode1 data-type Sectors. With CD-ROM/XA a track contains only Mode2 sectors. There are two form types within a Mode2 sector: Form1 contains user data (2048 bytes) and Error-Detection Code (EDC) and Error Correction Code (ECC) data. Form2 contains raw data (2324 bytes) such as audio or voice. For a CD-ROM drive to be fully XA comparable, it must have the following capabilities:

- Read data from two differently defined data streams. These are Mode2, Form1, which is static information, and Mode 2, Form 2, which is time-dependent information.
- Allow data from each stream to be buffered and delivered to the Central Processing Unit (CPU) and video subsystem as required.
- Translate the Adaptive Differential Pulse Code Modulation (ADPCM). It is a standard for audio compression, usually a 4:1 ratio. In other words, the CD-ROM drive must be able to send the audio signal to the speakers properly decoded and decompressed.

Nearly every CD drive you will encounter today is CD-ROM/XA compatible to the point where this is a nonissue. Occasionally you may run into an old CD drive that is noncompliant: but never a new one.

3.4.3.5 MULTISESSION CAPABILITY

Recall from the earlier discussion of CD-R that you can write to a CD-R only once. Multisession capability provides a way out of that limitation, but at a price.

You have a 640MB capacity CD-R and you write your daily backup of your data files to it. That takes up 100MB. The next day, you want to back up your data files again. You can create a second session on the same disc and write the next day's backup to the new session. The only gotcha is that once you create the second session, the first session is gone forever. In most cases, you cannot access multiple sessions, only the last one written to the CD-R.

A multisession CD-R drive has the capability to write multiple sessions. A multisession CD-ROM has the capability to read a disc that contains multiple sessions. (That is, it can read the last session written to the disc.) The most important reason an average consumer would need multisession support is to read photo CDs (such as a CD you would get back with developed film). All modern CD-ROM drives are multisession capable, as are all DVD and CD-R/RW drives.

3.5 CD / DVD Drive Connection

You can connect CD drives to the rest of your computer via followings

- Enhanced Integrated Drive Electronics (EIDE)
- Serial ATA (SATA)
- Small Computer System Interface (SCSI)
- legacy parallel, Universal Serial Bus (USB)
- Fire Wire port

Although you should be familiar with these interfaces from earlier unit, we will briefly review them here.

Before we begin, however, we will point out a couple of facts about CD drive interfaces.

First, they work in pretty much the same way as halted disk interfaces or floppy interfaces, so if you have a handle on those, you are set.

Second, most interfaces will do a credible job (with the exception of the legacy parallel port, which is extremely slow and should be avoided at all costs for external drives-if you can even find an external parallel drive anymore).

3.5.1 EIDE

The same EIDE interface that runs a system's EIDE hard disks will also run CD and DVD drives on most systems. I say most because some old systems might not support an EIDE CD drive. For it to work, the EIDE interface on the motherboard must support the AT Attachment Packet Interface (ATAPI) standard, which was introduced with ATA-3.

Because EIDE interfaces are already on all PC motherboards, many CD-ROM drives use EIDE. EIDE CD and DVD drives work fairly well and are less expensive than other types. However, when two EIDE devices share the same cable, latency problems (that is, delays) can result that may cause a high-speed CD-R or CD-RW drive to have errors during writing.

3.5.2 SATA

SATA is the bright star in high performance drive interfaces these days, and it is expected to replace EIDE even in low-end systems in the next couple of years. When that happens, most CD drives will start coming in SATA models. For now, you should be able to use a converter to get an EIDE CD drive to run on a SATA interface. It will not be any faster, but if you are trying to get rid of all the old ribbon cables, that might help.

3.5.3 SCSI

SCSI is a competing standard to EIDE for connecting devices to a PC. SCSI has a reputation for being faster than EIDE in general. However, the gap is closing and CD drives run slower than hard disks on any interface anyway, so increased speed alone is not a compelling reason to choose SCSI over EIDE for a CD drive. One thing SCSI does do exceptionally well, however, is manage the sharing an interface among multiple devices. You could have several SCSI drives on a single interface, and each could be doing its own thing simultaneously with very little delay. That is why SCSI is the professional's choice for high performance disk arrays.

A drawback of SCSI is that most motherboards do not support it directly. You must add a PCI based SCSI adapter and then connect your SCSI devices to it. This adds to the overall cost by \$50 to \$200 if you do not already have a SCSI adapter, which is not a problem for the high-end enthusiast, but could make a difference to Joe Consumer. Furthermore, SCSI CD drives in general are more expensive than their EIDE counterparts are.

3.5.4 USB

USB has become a popular interface in the past several years for external devices. It is fast and easy to connect. You can connect/disconnect USB devices without shutting down Windows.

The original USB specification to become popular was USB 1.1, which supported CD-R/CD-RW drives of only four. Newer systems provide USB 2.0 support, which is a much faster interface and supports the newest and fastest CD reading and writing speeds.

Choose USB if you want an easy to connect external CD drive, but be aware of the USB standard that the drive and the motherboard's USB interface supports. If you use a USB 2.0 CD drive on a motherboard with only USB 1.1 support, it will work but will default to USB 1.1 speed limitations.

3.5.5 IEEE 1394 (FireWire)

Fire Wire is a less common interface for CD drives than the previously mentioned ones, but Fire Wire drives do exist. FireWire is a competitor to USB; it is a high-speed interface for connecting external devices. Its capability is roughly equivalent to that of USB 2.0. The main drawback to Fire Wire is that most motherboards do not support it natively, unlike USB, so you must use an add-on PCI expansion board to insert the needed ports. One benefit of a Fire Wire drive is that you can use it with either PC or Mac.

3.6 Characteristics

All CD drives look alike (within the limits of being internal or external drives), but looks are not the measure of performance. You can buy drives with various spinning speeds, access/seek times, data transfer rates, buffering techniques, and so forth. All these features directly affect drive predominance.

3.6.1 DATA TRANSFER RATE

Data transfer rate, measured in kilobytes per second (Kbps), is the theoretical maximum amount of data that can be transferred to the PC per second. It is measured not in the actual kilobytes per second, but in "X" ratings. The original CD drives were 1, which meant they could read seventy five 2048-byte sectors of data per second. That comes to 150KBps. The next generation of drives was 1X, which doubled the speed by spinning the disc twice as fast past the read head. You can determine a drive's data transfer rate by multiplying its rating by 150KBps. In the following table, we have shown various speeds of drive and with respect to their data transfer rate and access time.

The improvement of data transfer rate is tabulated below.

DRIVE SPEED	TRANSFER RATE (KBPS)	ACCESS TIME (IN MILLISECONDS)
1	150	400
2	300	300
3	450	200
4	600	150
5	900	150
8	1200	100
10	1500	100
12	1800	100
16	2400	90
18	2700	90
24	3600	90
32	4800	85
36	5400	75
40	6000	75
48	7200	75
52	7800	

In reality, the data transfer rate for a drive is seldom if ever achieved because that rating is for sustained performance. It does not take into account the Mass time, which is the time required for the disc to rotate to the correct spot to be read and the read head's arm to move to the correct position.

DVD drive speeds cannot be fairly compared to CD drive speeds using ratings alone because DVD data is stored so much more compactly than CD data. If a DVD drive were a 1 drive, it would have a theoretical maximum transfer rate of about 1.4 Mbps, which is equivalent to about 9 for a CD drive.

It is not completely accurate, however, to simply multiply the X rating for a DVD drive by 9 to compare it to other drives because a DVD drive will also be used to read regular CDs. When evaluating the ability of a DVD drive to read CDs, you can approximately triple the rating to compare it to a CD drive. In other words, a 16X DVD drive would roughly compare to a 48X CD drive when reading from regular CDs.

Attention: The speed of the DVD-ROM drive does not affect the quality of a DVD movie being played. For example, if you have a 2X DVD-ROM drive, replacing it with a 16X DVD-ROM drive would not improve the quality of the movie playback. The speed will influence some DVD-based interactive multimedia games and browser.

3.6.2 ACCESS TIME

Access time is the amount of delay between the drive receiving the command to read and the actual beginning of the reading. It is measured in milliseconds (ms). The measurement is just an average; the actual speed depends on where the data is located on the disc and how quickly the read mechanism can get to it. The closer to the center the data is, the quicker it can be accessed.

Do not expect an access time from your CD drive that even approaches the access time you see from your hard disk. We have a hard disk with an access time of 9ms, but a common CD drive access time is more than 100ms. In above table access times are to be shown.

Notice that unlike data transfer rate, the access time does not get better at a constant rate as the drive's X rating goes up. The access time is a function of the quality of the hardware inside the drive that moves the read/write head. It is not directly related to the rotational speed of the disc, which is what the X speed measures. The relationship between them is tangential-the faster the drive speed, the more likely it is to have modern, high-quality head movement mechanism.

3.6.3 CACHE/BUFFER

Disk caching also called disk buffering, temporarily stores recently accessed or frequently accessed data to the hard disk to take advantage of its higher access rates. Most CD-ROM drives have a small amount of memory in them for this purpose. Typically, the directory of the CD is cached. Caching the directory enables the computer to more quickly navigate subdirectories and makes the CD-ROM drive appear to be faster. However, the actual reading of the data is still slower.

3.7 Installing a CD or DVD Drive

The installation procedure for a CD or DVD drive depends on the interface you have chosen. EIDE is different from SCSI, which differs from USB, which differs from SATA and so on. The following sections outline the basic procedures for each interface, with some detours that apply to multiple types. In the following figure we have shown the CD or DVD drives.



Figure: CD or DVD drives

3.7.1 Preparing a Drive Bay

Preparing a drive bay applies to all internal drives. Decide which drive bay you will place the drive in and clear the way for it. That may involve removing a bezel or cover plate on the front of the PC, moving any cables out of the way, and so on.

3.7.2 Setting Jumpers

Setting jumpers applies to EIDE and SCSI drives. You must set jumpers on the drive to let the system know what position or role this drive fills.

For EIDE drives, you must set a jumper to the Master, Slave, or CS position. Each cable must have only one master drive on it, so if you are installing a drive on an existing EIDE cable, you must set the second drive to Slave.

For SCSI drives, you must set a jumper indicating the SCSI 10 number. Each device in a SCSI chain must have a unique 10 number. You may also need to set a jumper for termination to indicate whether this drive is the last SCSI device on the chain.

3.7.3 Installing the Drive in the Bay

Again, installing the drive in the bay is for internal drives only (EIDE and SCSI). Insert the drive in the bay so that its front aligns with the front of the PC; then tighten it down with screws. On some PCs, you do not use screws to attach drives; instead, you use mounting rails with clips on them that snap into place. Some people prefer to attach the cables before tightening the drive into place.

3.7.4 Connecting the Cables

For an internal drive, connect a power plug from the PC's power supply to the drive. CD drives use standard Molex power connectors, which are the ones with four wires and four rather large round holes on them that plug into the four protruding cylindrical pins on the back of the drive.

Also for an internal drive, connect the ribbon cable from the EIDE or SCSI interface to the drive. The cable has a red stripe on one side; the red stripe goes toward pin1. If you look closely at the drive, you will see a little 1 at one end of the ribbon cable connector; that is the pin1 end. If you cannot find just orient the red stripe at the end closest to the power plug.

If you plan to play audio CDs on an internal CD drive, you will need to connect the drive to the sound card with an audio cable. This is typically a thin round cable consisting of three or four thin wires, with a small plastic three or four-pin connector at each end.

Okay, here is the first step for the external drives: connect its data cable to the PC using whatever interface is appropriate (SCSI, FireWire, or USB). Then connect its power cord. An external drive has its own power cord that plugs directly into an outlet.

3.7.5 BIOS Configuration

For a USB, FireWire, or SCSI drive, you do not need to do any BIOS configuration. (BIOS stand for Basic Input/Output System). The operating system will handle the drive automatically. Some SCSI controllers have their own BIOS Setup routines you can access.

For an EIDE drive, on a newer system you should not have to do any BIOS configuration either. The BIOS should auto detect the new EIDE device at start up, as your system boots, you will probably see a list of detected hardware; if your new CD or DVD drive appears on that list, the BIOS has figured it out.

Even if you do not see anything indicating that the BIOS has detected the drive, go ahead and let your operating system load. If the new drive appears there. You are fine-no worries man. On some old systems, you must enter the BIOS Setup and set the drive type for the EIDE interface to Auto for auto detection or to CD-ROM or ATAPI.

3.7.6 Drivers

Windows will detect and load the appropriate basic drivers for most CD and DVD drives automatically. These are protected-mode drivers. If you boot the PC to MS-DOS or boot from a startup floppy, however, you will not have access to the protected mode drivers, so you will need a real-mode driver for the CD drive to use it from a command prompt. The startup floppies that you can create in Windows 98 and Windows Me or XP include generic CD drivers that will work with almost all CD and DVD drives. Windows 95 startup disks do not have them. However, you can add them manually to a Windows 95 startup disk. Windows NT, 2000, and XP do not allow a command line boot from a floppy, so it is not an issue with those operating systems.

To operate a CD drive in MS-DOS or a command line outside of Windows (as in the case with a boot disk via Windows 95), you need to load the following drivers:

- If it is a SCSI CD drive, there needs to be a line in CONFIG.SYS that loads the driver for the SCSI adapter. This driver comes from the SCSI card's manufacturer. It might look like this: `DEVICE=C:\SCSI\SCSIDRV.SYS`
- In CONFIG.SYS, there needs to be a line that installs the driver for your specific model of CD drive. This driver comes from the CD drive's manufacturer. It might look something like this:
`DEVICE=C:\MYCD\NECCD.SYS /D:mscd001`
- In AUTOEXEC.BAT, the Microsoft CD extensions need to be loaded (MSCDEX.EXE). That would be something like this:
`C:\DOS\MSCDEX.EXE /D: mscd001` if loading from MS-DOS

If you need a driver for a specific SCSI card or CD-ROM drive, check the manufacturer's Web site. Generic drivers are also available that might work. You see them at work on a startup disk created with Windows98 and Me/XP, for example; they enable you to start your system and use your CD-ROM drive when you are not in Windows. In a pinch, you can copy those drivers to your hard disk and use them for running in MS-DOS mode.

3.8 Troubleshooting of Drives

Having problems getting a newly installed CD or DVD drive to work? Here is a step-by-step procedure for troubleshooting:

1. If it is an EIDE drive, do the BIOS see it? Go into BIOS Setup and check. If not, check for proper cable orientation, connection, and proper master/slave jumper settings. Also, make sure the EIDE channel in the BIOS is not set to Disabled. It should be set to CD-ROM, ATAPI, or Auto.

2. If it is a SCSI drive and Windows does not see it, check for proper SCSI termination and ID number.
3. If the BIOS see the drive but Windows does not, make sure that Windows is not in Safe mode. Windows will not see CD drives in Safe mode. Reboot if needed.
4. If the drive is not recognized when booting to a command prompt make sure the CD driver is loaded in CONFIG.SYS and that MSCDEX.EXE is loaded in AUTOEXEC.BAT
5. If data CDs will play, but not music CDs, make sure you have connected an audio cable between the drive and the sound card. For external CD drives, consult the documentation to find out whether an audio cable is required.
6. If a certain CD would not work, make sure the CD is clean. Clean it with a very soft cloth. Do not scrub. If you must use liquid, use alcohol. Also, make sure the CD is in a format supported by your current operating system; for example, make sure you are not trying to read it Mac CD on a PC.
7. If data DVDs will play but movie DVDs will not, make sure you will have MPEG2 decoding capability. This can either be hardware or be a piece of software. If it is hardware, make sure the cabling for it is correct per the instructions. On some boards, you must connect the monitor to the MPEG2 decoder board, then connect that board to your video card, and finally run a cable from the DVD drive to the board. Also, make sure that the correct driver's are installed and that your DVD player software is compatible with your model of DVD drive.

3.8.1 Recording CDs and DVDs

Windows 95 and higher will automatically recognize CD and DVD drives, but it will see them as ordinary CD-ROM drives. If you want to access their special capabilities, you must use third-party software. (Windows XP is the exception-it recognizes and uses writeable most CD and DVD drives automatically through its own software.) You can use the software that came with your drive or some other application.

3.8.2 Avoiding Buffer under runs:

One of the most frustrating parts about making a CD-R is encountering an error called a buffer under run. This occurs when the drive you are reading from cannot keep up with the CD writer's need for speed and makes it wait. The problem is that it cannot wait. It has to keep moving, and writing, at a more or less continuous rate. If it is ready to write but the data has not arrived yet, a buffer under run occurs, and you have just wasted a blank CD.

With newer CD recorders, this is less often a problem, they have large built-in data buffers designed to prevent this. With an older system or CD burner, the following can cause the dreaded buffer under run:

- Some other application is hogging the processor's attention, so the processor cannot tell the source drive to send the data to the CD-R fast enough. Try not to use another programs, or even move the mouse around, while a CD is being created.
- The source drive cannot operate as quickly as it needs to because of its own limitation. In that case, try transferring the files for the CD to your hard disk first and then making the CD from there instead of going from CD to CD.
- You do not have a large enough data buffer set up. A data buffer holds information read from the data source so that even if there is a pause in the data reading, the writer can keep writing without a wait. You typically set this up in the drive properties.
- You do not have: Direct Memory Access (DMA) enabled for the source drive. DMA modes transfer data with less CPU intervention. Try displaying the drive properties and turning on DMA transfer if possible.
- You are trying to record at too high a speed. Many CD-R drives advertise that they can record at up to 24X speed, but if you find you get buffer under run errors at that speed, try a slower speed such as 8X, 6X or even 4X, and your errors may go away.

In general, SCSI CD-R drives are less susceptible to buffer under runs than EIDE models.

3.8.3 Recording CD and Data DVDs

For CDs and data DVDs, if you are using third party writing burning software such as Roxio Easy CD Creator or Nero Burning ROM, you can choose to make an audio or data CD or DVD through that interface, consult the software documentation.

If you are using Windows XP's built-in writing tools, the procedure for creating data versus audio CDs is different. To do an audio CD, you go through Windows Media Player; for a data CD, you use Windows Explorer to drag and drop the files to the CD and then write them.

Attention: If you plan to use the CD or DVD on a computer with a different operating system, make sure the files have names that are recognizable in that operating system. For example, MS-DOS filenames are limited to 8 characters.

3.8.4 Producing and Recording DVDs

Developing DVD content is different from creating a CD, depending on whether you want to produce a DVD-ROM or a DVD-video

Full-quality DVD-video DVDs, however, are another matter entirely. You develop a DVD-video in three stages. First, you encode the video. Then you design, configure, and test the interface. Finally, you premaster a disc image.

If you want to go the distance, you can put your home movies onto DVD-video by following these steps:

1. Use video and audio capture boards to capture your video and audio from VHS, Hi8, or DV. You can use service bureaus or scanners to transfer slides to digital format.
2. Use a software encoder to encode the video into MPEG2 format. Set the video frame rates to 29.97 frames per second (fps) for NTSC format (U.S.) or 25fps for PAL (European) format.
3. Use a software or hardware encoder to encode the audio into Dolby Digital. Format the audio as 48 kilohertz (kHz) Pulse Coded Modulation (PCM).
4. Import the video and audio clips into a DVD-video authoring program. Create menus and buttons that link to your media clips. You can also import your slides in TIF, JPEG, or Photoshop format.
5. Write your completed project to a DVD-R.

3.8.5 Selecting Writeable Media

When selecting writeable CDs or DVDs, you have a wide range of choices. One blank is distinguished from another in the following ways:

Standard supported: Make sure you get the right type of blanks for you drive and what you intend to do with it CD-R, CD-RW, DVD-R, DVD+R, DVD-RW, DVD-RAM and so on.

Rated maximum write speed: High quality discs have higher maximum write speeds. If your CD-RW drive maxes out at 8X write speed for CD-R, there is no reason to spend extra to buy 32X blanks. Match the blanks to the speed at which you intend to record.

Capacity: Standard CD-R and CD-RW blanks hold 640MB of data; high capacity ones hold 700MB. Very old writeable drives might not be able to use the high-capacity ones.

Dye and metal colors: It can be difficult to know what you are getting in this area until you open the package because most labels do not say. The standard CD-R uses green dye and gold metal; these CDs work well in most players. Gold dye/gold metal is a more sensitive combination that might not play well in an audio CD player. Blue dye and silver metal produces a longer-lasting CD, up to 100 years.

3.8.6 Handling CDs and DVDs

By caring for your CDs and DVDs, you are doing preventive maintenance on the drive, reducing the opportunity for contaminants to enter the drive. The following pointers will keep your CD and DVD drives up and running:

- Handle the CDs and DVDs only at the hub or the outer edge. Do not touch the shiny surface.
- When you insert a CD or DVD into the drive, make sure you seat it properly in the tray.
- If a CD or DVD needs to be cleaned, use a soft dry cloth and wipe in the radial motion from inner hub. Do not use a circular motion.
- Do not use cleaning agent. Because solvent can damaged the surface of the discs
- Do not use wet clothes
- Avoid cleaning label side of disc.
- Use caddy to transport CD or DVD and store in cases when not in use.
- Avoid exposure to heat or cold.
- Avoid direct sunlight and Ultra violet light.

3.9 Summary

In this unit, we have discussed about CD / DVD and their drives. We have also described installation and configuration of these devices. We have introduced different components of CD and DVD drives. Finally, we have explained about trouble shootings and maintenance of CD, DVD and their drives.

3.10 Model Question:

Write a short note on the following.

1. What are roles of drives? Can you explain CD, DVD in brief?

2. Give the comparative picture of CD and DVD in computer field.

3. Write the installation and maintenance CD and DVD drives. Describe their advantage over floppy disk drives.

UNIT- 4: INSTALLATION AND MAINTENANCE OF FLOPPY DISK DRIVE

Structure:

- 4.0 Objectives**
- 4.1 Introduction**
- 4.2 Removing Old Floppy Drive**
- 4.3 Obsolescence of Floppy Technology**
- 4.4 Alternatives of Floppy Drive**
- 4.5 Floppy Subsystem**
- 4.6 Floppy Disk**
- 4.7 Floppy Drive**
- 4.8 Disk Controller**
- 4.9 Connecting Cable**
- 4.10 Maintenance of Floppy Disk**
 - 4.10.1 Heads Cleaning**
 - 4.10.2 Defending Disks**
- 4.11. Configuring and Installing Floppy Drives**
 - 4.11.1 Removing Floppy Drives**
 - 4.11.1 Installation Floppy Drives**
- 4.12 General Troubleshooting**
 - 4.12.1 A Disk Cannot Be Read**
 - 4.12.2 Drive Refuses the Function**
 - 4.12.3 Drive Shows Phantom Directories**
- 4.13 Special Troubleshooting**
- 4.14 Summary**
- 4.15 Model Question**

4.0 Objectives

At end of this unit, student will be able to:

- Understand about the basic function of floppy disk and drive.
- Explain the various kinds of floppy disk and drive its components.
- Discuss the installation of floppy disk and drive.
- Know the concept of various types of interfacing devices for floppy disk and drive.
- Analyse statistics of various capacity of floppy disk.
- Generate awareness about maintenance of floppy disk and drive.
- Identify various trouble shootings for floppy disk and drive.

4.1 Introduction

IBM originally created floppy disk drive in 1967; the reference is derived from the floppy media that is encased within the protective casing. This media is a magnetic medium and is written to much like the method used by hard disk drives. Floppy drives have been found in computers for several years and are commonly still being used today. The next generation of floppy drives will more than likely be the LS-120 diskette drives, which are already being included in many computers now.

The floppy disk is a very old technology, preceding even the hard disk. There have been various types over the years, each new technology improving the performance and reliability levels, but we are still talking about really old stuff here.

So why have floppy disks hung on this long? Mainly because until recent years there was not any other good, way to move files between computers that everyone could agree on. Many manufacturers have come up with better ideas over the years for floppy like storage, but none has become industry standards. Finally, in the past several years, some compelling floppy alternatives have become in mainstream. and computer manufacturers are finally phasing out the old floppy technology. In this unit, I will explain the basics of floppy technology and describe some modern alternatives to them, we will also explain how to install a floppy disk drive and how to troubleshoot floppy operation.

4.2 Removing Old Floppy Drive

First, let us take a quick look at the procedures and guidelines for replacing your floppy drive. Before start the operation, you will need followings tools.

- Nonmagnetic Phillips-head screwdriver
- Container to hold the removed screws
- Antistatic wrist strap
- New floppy drive

And, as usual before you work on a PC, ensure that the PC is turned off and unplugged.

Attention: Floppy drives are not especially sensitive to static electricity, but installing one requires connecting it to the Input/Output (I/O) controller with a ribbon cable; and the motherboard can easily be harmed by Electrostatic Discharge (ESD). Wear that antistatic wrist strap.

Hint: Before replacing a floppy drive, make sure the problem is not with an individual disk; Try the drive with several disks.

To replace a floppy drive, follow these steps:

1. To replace a floppy drive, first turn off the PC's power and remove the cover.
2. Disconnect the ribbon cable and the power connector from the drive. Notice which direction the red stripe on the ribbon cable points (toward the center or the outside of the drive).
3. Unscrew the screws holding the floppy drive in place. Pull the drive out of the PC.
4. Install the replacement drive by using screws or clips, as the case requires.
5. Connect the power cable.
6. Connect the ribbon cable with the red stripe going to pin1. The drive may have a little notch next to one end of the connector to tip you off; if not, connect it in the same way it was connected on the old drive (see step 3).
7. Turn on the PC and test the new drive by trying to read a disk from it.

4.3 Obsolescence of Floppy Technology

Floppy drives are well on their way to being obsolete. Floppies used to be irregular in a system because they were the only removable disk type available. If you needed to transfer a file from one PC to another, a floppy was your only option. Today, however, there are writeable CDs, Universal Serial Bus Flash devices, Zip disks, the Internet, and Local Area Networks (LANs) are all of which do a better job than a floppy disk and can hold more data.

However, do not expect the PC industry to let go of them just yet. New computers will probably continue to come with floppy drives for at least another couple of years, until the average of consumer realizes that it is really does not need one after all.

One reason that many people hang onto having a floppy drive is that they think they need to be able to boot from a startup floppy. In the event if a system has problem. And this is a valid point for people who are running 9x versions of Windows. However, the NT based Windows versions all come on bootable CDs that kick into a Setup utility with a Repair option, so for modern Windows versions a bootable floppy is not an issue. In addition, you can use a CD burning application such as Roxio Easy CD Creator or Nero Burning ROM to make a bootable CD-R if your own that contains an image file of a startup floppy. When the PC boots from it, it

simulates a floppy disk. Therefore, the actual physical floppy disk drive is not a requirement.

4.4 Alternatives of Floppy Drive

You can probably name at least a half-dozen better ways to move data than a floppy disk, right. We will describe a few of my favorites before getting into floppy technology in details:

Flash Random Access Memory (RAM) devices: This is the newest, coolest thing. It is a little disk/drive combo that connects to the PC via a USB or Fire Wire port, and the computer sees it as a drive. When you are done with it, just disconnect it and plug it into different computer instant portability.

Note: There are many different kinds of Flash RAM devices because this technology is in the innovation phase where manufacturers are getting creative trying to capture the public's interest in the war to become the established standard.

CD-R: It is one time, writeable CDs-cheap, plentiful and readable in almost every CD drive. These hold about 700 megabytes (MB) of data.

CD-RW: These CD blanks are rewriteable up to, 1000 times with packet-writing technology and hold about 700MB of data.

Attention: The Operating system must be able to read Universal Data Format (UDF) to read a CD-RW or you have to install a UDF reader driver in the operating system. UDF is a packet-writing technology driver that allows the operating system to read the special type of data storage that CD-RWs use. Windows XP comes with it; so does the CD-RW drive's writing software.

LS-120: This is also called Super Disk. It is a 120MB floppy disk, which looks just like a normal floppy disk but is readable and writeable only in LS-120 drives. LS-120 drives can also read standard floppies. This was a good idea but never really caught on, and now there are better alternatives.

Attention: Even though an LS-120 drive looks almost exactly like a regular floppy drive, it runs on the Integrated Drive Electronics (IDE) interface, not the floppy interface. Therefore, if you are setting up a system, keep in mind that going with an LS-120 drive as a floppy replacement will eat up one of the four Enhanced Integrated Drive Electronics (EIDE) positions on the motherboard.

ZIP: Once the up-and-coming darling of portable storage, I/O mega Zip disks have lately taken a back seat because the popularity of writeable CD. A Zip disk is a cartridge that is slightly larger than a floppy disk; it fits into a Zip drive and holds 100MB, 250MB or 750MB of data, depending on the model. Zip drives are either SCSI or IDE. The original Zip drive also came in an external parallel model, which is now obsolete.

JAZ: I/O mega used to make a larger disk drive called Jaz, that held 1 or 2 gigabytes (GB) of data, but they no longer do. You may still see such a drive in the field, however, and blank disks for it are still sold.

4.5 Floppy Subsystem

Now, we are done trashing the floppy drive interface, for now anyway. We will assume you are still working with it and want to know something about it. We will explain the mechanisms behind a floppy drive and its disks and tell you a little about troubleshooting and repairing floppy drives should you decide that it is worth your while to do so.

The floppy disk subsystem like the rest of the PC is modular. That is useful because you will see that you can use this modularity to **divide and conquer** in order to solve problems. The subsystem consists of four parts:

- Floppy disks
- Drives
- Disk controllers
- Connecting cables

4.6 Floppy Disks and Their types

A floppy, when extracted from its case, looks like a thin 45-rpm record, only quite a bit smaller. It is a Mylar disk with iron oxide affixed to it.

We worry so much about dust getting on our hard disk surfaces that we seal up the disk drive. Should we have the same concern about floppies? Is there some way to clean a disk? Those of us older than, say, 30 may recall getting out the Disc washer and cleaning an album prior to playing it. For some people, a ritual such as this was comforting. For those people, I must sadly report that no such ritual occurs with floppies.

Floppies are stored inside their own Discwasher: a semi rigid case lined with flexy material. As the disk rotates inside the case, the material picks up any dust. The case has a hole cut in it so that the disk can be read/ written without having to remove it from the case. In general, there is no need to clean a floppy disk.

A reasonable life expectancy for a floppy disk, according to disk manufacturers, is about three to four years. There are, of course, better and worse disks. There are different coating; on higher-density disks, generally incorporating less iron oxide (a cheap ingredient) and more cobalt or barium (expensive ingredients). Floppy Disks are cheap enough these days that they are disposable; if a disk comes up with an error, just throw it away.

Attention: You should not archive data on a floppy disk that you expect to need later several years because floppies tend to develop surface errors over time that make certain sectors (it is a crapshoot which ones) on the disk unreadable. Try to store important archives: on CD-R or at least Zip disks.

Different sizes and capacities of floppy disks are used, but nowadays there is only one only one of high-density, double-sided floppy with a formatted capacity of

1.44MB disk is popular in the field. Occasionally, you may run into other capacities and sizes.

8" Disk - The first disk was introduced in 1971. The disk was 8" in diameter with a magnetic coating, enclosed in a cardboard case with the capacity of one megabyte. Conversely, to hard disks, the heads actually touch the disk, like in a cassette or video player, which actually wears the media.

5.25" Disk - 160KB first Created in 1976 and later becoming a standard in 1978.

Dimension	Structure	Capacity
5.25"	Single Side	160KB
5.25"	Double Side	360KB
5.25"	Double Side High Density	1.2MB

3.5" Disk -IBM created it in 1984, which even today are still commonly used. Most 3.5" newer disks have a capacity of up to 1.44MB.

Dimension	Structure	Capacity
3.5"	Double Density	720KB
3.5"	High Density	1.44MB
3.5"	Extended Density (IBM ONLY)	2.88MB

3.5" FLOPPY DISKETTE

Below we have represented a picture of the now commonly used 3.5" floppy diskettes. The below representation is what the back of a floppy diskette looks like. As you can see, there are four main visible parts:

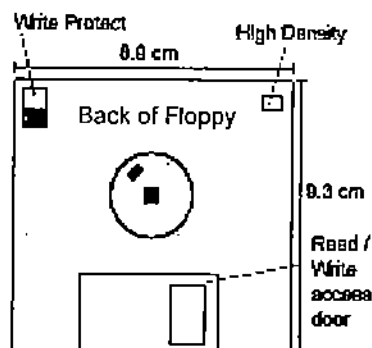


Figure: 3.5" Floppy Disk

First, the top left of the back of the floppy there is a small tab. This tab enables the floppy disk to be switched from write protected and un-write protected. Move the tab to the top position, creating a hole and make the disk write protected. This means that nothing can be written, erased or deleted from the diskette. Moving the tab to the bottom position allows the disk to be un-write protected, which means the

diskette can be written too, erased and have information deleted from the diskette. Some diskettes, which are generally cheaper diskettes, will be missing this tab. To write information to the diskette you will have to place a piece of scotch tape over the hole.

Second, the top right hand of the back of the floppy you will notice a small hole. This tells the computer if the diskette within the computer is a High Density diskette.

Third, you will notice the circular metal disk in the middle of the floppy diskette. This is used to rotate the magnetic medium within the floppy disk casing.

Fourth, you will see a metal door, which can be moved left and then will snap back to its original position. This door is used to allow the read/write head within the floppy drive to have the capability of accessing the magnetic medium within the casing. Once the diskette is removed, this door will snap back into position, helping to prevent anything from getting on the magnetic medium and destroying it.

5.25" FLOPPY DISKETTE

Below is a representation of a 5.25" floppy disk, commonly used in 1980's and is generally no longer used today. The 5.25" floppy diskette was floppy when compared to the 3.5" floppy diskette.

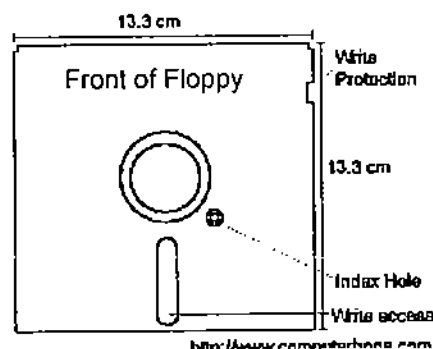


Figure: 5.25" Floppy Disk

As you can notice in the above illustration, the 5.25" floppy does not have many involved components.

First, you will notice the Write protection tab, which would be present if the diskette allowed information to be written to it. Placing a piece of tape over the hole or if the hole was not present would make the diskette write protected.

Second, you should notice the large hole in the center of the floppy diskette. This hole allowed the platter within the floppy to be rotated allowing the information to be read from the write access hole.

Third, you should notice the Index hole, this hole was the computer's way of knowing where the starting point of the sector was on the diskette by physically having a hole punched in the film disc within the plastic casing.

Finally, you should notice the write access hole. This is the location which information would be read from as the disc within would rotate. However, because there was no **protection** like that found on the 3.5" floppy diskette, the 5 1/4" floppy diskette was generally protected by placing it within a protective sleeve.

LS-120 looks and acts like a floppy disk, but technically it is not exactly a floppy disk. It uses a combination of magnetic and optical technology. The LS stands for Laser Servo, referring to the fact that it uses a laser for alignment in the drive. Then it writes the data using magnetic data tracks, much as an ordinary floppy does.

4.7 Floppy Drive

For many years now, 3.5-inch, 1.44MB floppy drives have been the standard; you did be hard-pressed to find one of the old 1.2MB 5.25 inch drives anymore except on the oldest PCs (we are talking XT here). Even if you did find such a drive, finding a floppy disk for it that has not gone bad would be an even greater challenge. Remember, the average life of a floppy is only a few years.

The first 3.5-inch floppy drives were not KB in capacity, rather than 1.44MB, but again, such drives will appear only on very old systems. The original not KB floppy disk drive originated with the IBM PC-AT, which was an 80286 system produced in the mid-to-late 1980s. Therefore, that gives you some idea of the age of the thing.

A 3.5-inch drive connects to the drive controller (normally built right into the motherboard) with a ribbon cable, the same as a hard drive or CD-ROM drive (except the ribbon cable the floppy drive uses is slightly narrower-J4 wires instead of the 40 used for EIDE). It also requires power. Power Supplies and Power Protection that a floppy disk uses the smaller Berg connector (also called a mini) from the power supply and not the normally chunky Molex connector.

Hint: Most power supplies come with only one Berg connector, so if you have two floppy drives, you will need a power cable adapter. Your computer's motherboard or power supply, the case may come with such an adapter; if not, you can pick up an inexpensive one at a computer store. However, it is rare that anyone would have two floppy drives anymore as we said earlier; even one floppy drive is arguably too many in a modern system.

4.8 Disk Controller

Whenever a computer wants to interface with an outside device, it needs a controller to act as a go between to allow the outside device-a floppy, in this case to talk with the Central Processing Unit (CPU). This functionality is handled through the motherboard's chipset (the Super I/O portion of the chipset, to be exact). A single floppy controller can handle up to two floppy drives on a single ribbon cable.

4.9 Connecting Cable

The last piece in the floppy subsystem is the cable that connects the controller to the drive. Cables sound insignificant, but this one is not: I have fixed many floppy problems by swapping out defective cables. The drive is connected to the controller by a 34-wire ribbon cable. Most of the cables have three connectors: one for the drive controller, one for drive A: and one for drive B: (although few computers nowadays have a second floppy drive).

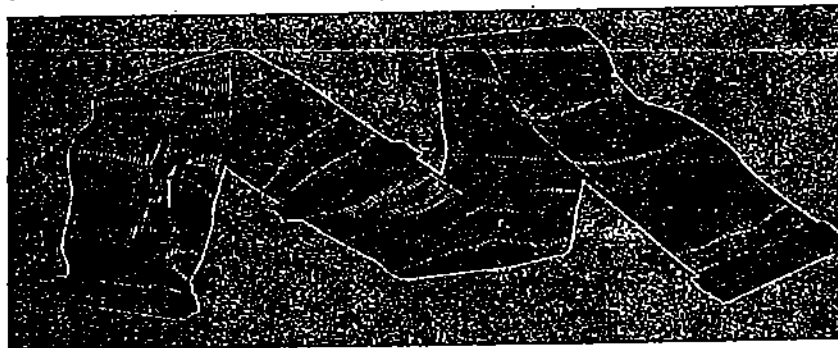
Attention: Because the cable is 34 wires wide, you would think that the interface between the floppy and the controller is a parallel interface 8 bits wide, but it is not. The floppy to control interface is serial, transferring 1 bit at a time.

It is important that you plug your floppy drive into the right connector on the cable. The connector that is farthest away from the other two is the one that goes to the motherboard (or I/O controller board). The one at the opposite end from that goes to the A: drive (the primary floppy drive). The middle one is optional and is used if you have a secondary floppy drive.

Why, positioning is more important? In the cable, there is a twist in part of the wires at one end of the cable. That twist changes the flow of the data between the B: drive (middle) and the A: drive (the twisted end), so the controller can distinguish between them when addressing them.

It never hurts to have a couple of extra ribbon cables; sound. The cables are inexpensive and available from many mail-order houses or computer discounters, so it is easy to keep all extra one around. Later in this unit, I will tell you about some of the common problems with floppy drives, one of which is having a broken wire or two in the ribbon cable. It is common and easy to fix (just swap the cable out).

The figure has shown the connecting cable use to connect drive with motherboard.



4.10 Maintenance of Floppy Disk

The failure of a floppy disk or drive can be a scary, potentially disastrous thing if you have stored some critical data on a floppy disk. (That is not a very good idea, by the way. Never put your only copy of something important on a floppy). What can you do to lower the probability of a floppy disaster?

4.10.1 Heads Cleaning:

Anyone who has VCRs and rents videotapes on a regular basis probably knows about the need to clean the VCR's heads. Something like that can be done for the floppy as well, but experts disagree as to whether it is worthwhile to do so.

Like your VCR, floppy drives have an electromagnetic read/write head that does most of the drive's work. In the process of accessing the disk, the head may rub some of the coating off the disk, it seems reasonable to assume that head cleaning kits are good things.

When you cannot see something, you get superstitious about it. We cannot see the junk on floppy heads (or the data on magnetic disks for that matter), so we play it safe and purchase a floppy head cleaner to ward off floppy evil spirits. Once we have procured the cleaner, however, our first question is, "How often should we clean the floppy heads?"

The instruction on the head cleaning is done weekly. That does not seem to be a good piece of advice; because a floppy head cleaner in a drive could misalign a disk head. The floppy cleaner is a relatively thick piece of cotton shape a like a floppy. Worse yet, some floppy cleaners are abrasive they wear a way some of the floppy head with every use. (One has to look askance at this procedure. Scraping away some of the head to find a clean part seems somewhat like using sulfuric acid to clean one's teeth.)

Personally, (we) clean heads only when the drive fails, we have some computers on which we have never cleaned the heads and they work fine. Our recommendation: clean the heads only when you start experiencing read/write errors. (Of course, for only slightly more money than a good quality head cleaning system, you could have a completely new floppy drive so weigh your options.)

As we hinted before, be careful when choosing a floppy head cleaner. Do not buy an abrasive cleaner. Make sure it uses a cotton floppy and some cleaning fluid.

Attention: By the way, there is an issue related to cleaning the heads called demagnetizing them; it is done with some audio recording heads.

4.10.2 Defending Disks: Did you ever notice the do not cartoons on the back of a floppy disk jacket? Do not expose the disk to heat, cold, magnets, or dust, they warn. They lead you to believe that floppies are very fragile items: A little practical experience with disks teaches you that it is not true. Yes, disks must be taken care of. But you need not get crazy about it. Do not put them on the radiator or leave them on a shelf that gets three hours of direct sunlight every day. Do not store them under the roof leak or use them as coasters. Given the choice, store them upright, stack left and right rather than on top of each other. In addition, keep them away from magnets.

When it comes to temperature extremes, you have to be concerned about thermal shock. If your portable computer has been sitting in the back of the car in freezing temperatures overnight, bring it in and let it warm up before using it. Just a little heat expansion or contraction can temporarily realign your drives or make the motors respond a little differently. And condensation from rapidly warming items can build up on the electronics, and the moisture can cause electrical shorts. Extreme temperatures can damage the disks; vendors claim that disks should never be

stored in weather colder than 50°C or warmer than 125°C. Cold is not as much of a problem as heat.

Dust, smoke, and dirt can cause damage to the head or a disk. Everybody knows smoking is not good for you, but you may not know just how bad it is for your drives. If you are in a dusty environment, think about getting a power supply with a filtered fan. Running the air through a filter before pushing it through the system removes the vast majority of the dust particles.

Magnets-both the permanent and electromagnetic type can cause permanent loss of data on hard or floppy disks. Something we have yet to understand is the little plastic paper clip holder with the circular magnet near the opening. One day, you will put a floppy on top of the paper clip holder. It will be sad. On the other hand, you will use one of the paper clips, now magnetized by their proximity to the magnet to clip a document to a floppy.

Attention: Watch out for less obvious, magnetic sources, too. Unshielded stereo speakers, for example, have powerful magnetic pulls, and some older telephones have magnets in them, too.

4.11 Configuring and Installing Floppy Drives

You generally will not repair drives; you will replace them. That means it is most important to be able to know about followings.

4.11.1 Removing Floppy Drives

You can remove floppy drives in three steps:

1. Remove screws from the mounting brackets.
2. Remove the power connection cables.
3. Remove the data connection cables.

4.11.2 Installing Floppy Drives

Installing a floppy drive is just the reverse of removing one. To install a new floppy drive, you must follow these steps:

1. Mount it into a drive bay.
2. Connect it to the after-the-twist end of the ribbon cable if it is drive A: or to the center connector if it has drive B: You cannot have a drive B: in a system without a drive A: so your primary floppy should always be A: at the opposite end of the cable from the controller.
3. Attach a power connector. Again, you want the smaller power connector, for the floppy drive.

Attention: Most of the PCs are to be used in standard sized floppy drive. But some of proprietary systems (the Compaq Presario is particularly notorious in this regard) are to be used as non-standard floppy drive with offset positioning, that requires you to buy a replacement directly from the original manufacturer (and typically at two to

three times the cost of a generic replacement). So if you are replacing a floppy drive in someone else PC, it pays to examine the old drive before buying a replacement for it.

4.12 General Trouble shootings

The mechanical nature of floppy drives makes them prone to a host of ills. Few of them are listed below.

4.12.1 A Disk Cannot Be Read: Pop a disk in the drive, try to read it, and the message says Data error reading drive A, Sector not found reading drive A, or perhaps the dreaded General failure reading drive A. Such a message is even better than your morning coffee to get the old blood pumping, particularly if the disk contains your only copy of the football pool.

Do the following:

1. Press R for Retry, Sometimes, it will work. Ignore the painful grinding sounds. After three or four unsuccessful retries in a row, you may safely conclude that it is not going to work. However, if the drive seems to read a little bit of data before repeating the error, keep trying; you may be able to conclude the drive into reading the file-perhaps just this once before it goes to put completely.
2. Remove and replace the floppy disk. I have seen 3.5-inch drives that are unreadable until you take out the disk and put it back in.
3. Take out the floppy disk and try it in another drive, The read write had on one machine maybe aligned just slightly differently than on another, and that little difference might make one drive read a disk that another cannot. If one drive can read a disk that, another cannot, either the disk drive's head could be misaligned or the floppy disk could have been formatted by a different disk drive and that drive's head was misaligned. Try to find the drive that created the disk.
4. If you still have not solved the problem, you are in the same boat as someone who is suffered a media failure on the hard disk.

4.12.2 Drive Refuses the Function: When the drive will not read or write properly, there are bunches of possible causes. We are assuming that this is the drive which worked fine yesterday and to which you have not done anything that would obviously cause drive problems. Follow these steps:

Did you see a floppy drive controller failure error, on the screen when the system booted up? Such a Power-On Self Test (POST) error message means that drive A: did not respond. The system no longer realizes that you have a floppy drive. It means that your floppy controller has failed (in which case both floppies-A: and B:- would be dead), the floppy drive itself has failed or the cable is bad.

Hint: If you have two floppy disk drives and only A: is recognized at boot up, not B: it could be that the CMOS battery has died on your motherboard, resetting the BIOS Setup configuration back to factory default settings (one floppy). Try going into BIOS Setup and reinforcing it that you have two floppy drives.

1. If there was no POST error, try other disks in the drive. If only one disk gives the drive fits, the problem more likely lies in the disk.
2. Clean the disk drive heads. It is easy and takes only a minute.
3. Try to format a disk. If you can format a disk all right but that disk is, unreadable by other drives your drive head is probably misaligned. If it is misaligned, new floppy drives are about \$10 to \$20 today.
4. Finally, swap the relevant components: the controller, the cable, and the drive. Swap only one component at a time, and when swapping a component does not solve the problem, reinstall the original component and swap another one.

Do not overlook the cable. Cables can be nicked when you are installing boards, replacing drives or just removing the cover. How bad that is depends on which line is nicked. One cable on an old laptop that we once owned kept the change line from working; when we put in a different floppy, the file listing still showed the previous disk. It takes only a two-minute swap with another cable to find out for sure.

4.12.3 Drive Shows Phantom Directories: Imagine that you display a file listing of a floppy disk's contents. Then you remove the disk and put a different disk in the drive. You refresh the file listing, but you see the directory of the previous floppy, not the one in the drive now.

If you see this symptom on a computer, this is a red alert. Do not use the thing until you get the problem fixed. The reason: say you put a floppy in the drive and display the listing of files on it. Then you put a different floppy in the drive and write some data to the floppy. The PC writes data to the new floppy using the old floppy's directory. That means the newly written file is probably okay, but everything else on the floppy is trashed.

Where does this problem come from? It is something called the change line signal. It means that on all modern systems, the 34th wire in the cable (the one farthest away from the red stripe) is in charge of letting the drive controller know when a new floppy has been inserted and rereading its table of contents. If there is something wrong with that wire in the cable or with that pin on the drive connector, you get this phantom disk problem.

Suppose you find yourself change line problem-phantom directories during working. What do you do? Check the floppy cable. Try substituting a different floppy cable and see if that does not clear up the problem. If it does not, the problem is inside the drive itself.

4.13 Special Troubleshooting

Before you give up on a floppy drive, try these tips:

- Always try other disks in the drive before assuming the drive itself is bad.
- If the floppy drive light comes on and stays on all the time, you have installed the ribbon cable backward.
- If the floppy drive light does not come on at all, check the power connector.
- If you see a floppy drive controller error at startup, check the BIOS to ensure that it recognizes your floppy drive. If the CMOS battery fails, it may forget about all drives except the first Floppy drive (A :).

- If you switch disks but the PC still shows the previous contents of the drive, check for a damaged ribbon cable.
- Intermittent disk access errors often indicate a bad power supply in the system.
- Data errors reading the disk can mean a bad disk. You might also try cleaning the disk heads by using a disk-cleaning kit.
- Try to format a disk. If it works but that disk is then unreadable by other drives, the drive head is probably misaligned. Throw the drive away, and get a new one.
- If replacing the drive does not help, try replacing the cable. If that does not work, the problem is probably the disk controller. If it is built into the motherboard, you may need to disable it in the BIOS and install a separate I/O controller card to use instead.

4.14 Summary

In this unit, we have discussed about floppy disk and drive. We have also described installation and configuration of these devices. We have introduced different components of floppy disk and drive. Finally, we have explained about trouble shootings and maintenance of them.

4.15 Model Questions:

Write a short note on the following.

1. What are roles of floppy disk in the computer field? Can you explain its functions?

2. What is role disk drive in the computer? Explain the different component of floppy disk.

3. How to face the trouble arise in installation and configuration of floppy disk and its drive. Write the steps of head cleaning.



BLOCK

7

**Maintenance and Trouble-Shooting
Of Monitor**

Unit 1 Overviews of Monitor	5
Unit 2 Different Monitors and their Video Adapters	15
Unit 3 Interfacing and Upgrading the Monitor	37
Unit 4 Maintenance of Monitor and Its Trouble-Shooting	42

BLOCK INTRODUCTION

Now Monitor is a common output device by which we can capable to visualize the every activity of computer. Day by day, some of advancement in the computer field is going to improve the technology. However, as we know quality of monitor is available based on their video adapter currently seen in this field? Therefore, we have needed to study such devices in the field of computer.

Now in this block, we will discuss about monitor. We will also explain about installation and configuration of the device. Finally, we consider the description of maintenance and trouble shooting of it that gives the real boost to generating hardware engineer those must become aware about monitor in the computer field.

We will mainly focus on following point:

Concept of Installation of Monitor: Concept of installation monitor is more necessary, when candidates are wanted to become maintenance engineer in computer field must be aware about previously mentioned devices.

Concept of Configuration of Monitor: Concept of configuration of monitor is also necessary, when candidates are wanted to become maintenance engineer in computer field will require knowing fact of configuration of monitor.

Concept of Maintenance of Monitor: Concept of maintenance of monitor is necessary, when candidates are wanted to hold the knowledge to become computer maintenance engineer.

Concept of trouble shooting of monitor: Concept of trouble shooting of monitor is compulsory for candidates who are wanted to do hardware maintenance in computer field.

Unit wise description of block is given as below:

The first unit of this block is concentrated on basics of different type monitor, which create conceptual view of function of different monitor.

In the second unit of this block, we have devoted on interfacing of monitor. We further discussed about its maintenance.

In the third unit of this block, we have mainly discuss about upgrading and interfacing of monitor , which are also essential for computer professional.

In the fourth unit of this block, we have described about maintenance of monitor and its trouble shooting.

Suggestion for further reading

1. PC Upgrade & Maintenance Guide, Fifteen Edition, Mark Minasi
2. Modern All about Printer, Manhar Lotia, Pradeep Nair
3. Inside the PC, Eight Edition, Peter Norton

UNIT-1: OVERVIEW OF MONITOR

Structure:

- 1.0 Objectives
- 1.1 Introduction
- 1.2 Basics of Monitor
 - 1.2.1 Pixels
 - 1.2.2 Resolution
 - 1.2.3 Raster Graphics
 - 1.2.4 Interlacing
 - 1.2.5 Non-interlaced Display
 - 1.2.6 Frame Rate
 - 1.2.7 Video Bandwidth
 - 1.2.8 Synchronizing Signals
 - 1.2.9 Horizontal Scanning Frequency
 - 1.2.10 Vertical Scanning Frequency
- 1.3 Overview of Picture Tube
 - 1.3.1 Monochrome CRTS
 - 1.3.2 Color CRT
 - 1.3.3. Resolution
 - 1.3.4. Shadow Mask
 - 1.3.5. Slot Mask (Aperature Grille)
 - 1.3.6. Screen Curvature
 - 1.3.7 True Resolution
- 1.4 Summary
- 1.5 Model Question

1.0 OBJECTIVES

At end of this unit, student will be able to:

- Understand about the basics of monitor.
- Explain the various kinds of monitor and its components.
- Know the concept various feature of monitor.
- Analyse statistics of feature for suitability of picture tube of monitor.
- Generate awareness about resolution of monitor.
- Identify the various clue of paper handling.

1.1 INTRODUCTION

When we use a computer, we need some device to see the results of the work what we are doing. A computer monitor is used for this purpose. Monitor works very much like a TV, main difference between TV and monitor is a monitor does not contain any tuner circuit and resolution (number of column and lines it can display) of monitor is much higher as compared to a TV.

Like a TV, main component inside a monitor is picture tube or CRT (Cathode Ray Tube). It is an evacuated glass tube, roughly conical in shape, with a phosphor coating on the inside of the large screen end and an electron gun at the narrow end. The gun fires a narrow beam of electrons at the screen and when this beam strikes the phosphor coating, light is emitted.

1.2 BASIC TERMINOLOGY

Two techniques are basically followed for generating image on CRT display; they are Raster Scan and Vector Scan. The raster scan technique is the universally accepted one used in CRT displays of the PC's. In raster scan technique; the horizontal and vertical deflection signals are generated to move the beam in forward and backward direction across the screen like a raster. The retrace portion of raster scan pattern is suppressed (blanked) by reducing the intensity of electron beam during the time of retrace. The complete CRT screen can be considered to be made of hundreds (or even thousands) of dots, commonly known as pixels (picture elements). The electron beam can illuminate any pixel when it hits the pixel during the raster scan process or reducing the intensity of electron beam can blank it. On a standard monochrome monitor, the electron-beam scans the screen 50 times per second. The phosphor has persistence, which means it keeps radiating light after the electron beam has moved on. The scan rate and persistence together make the user see a steady, flicker-free image.

A monochrome monitor needs three signals to produce a display, horizontal synchronization (HSYNC) to tell it when a scan line is about to begin, vertical synchronization (VSYNC) to tell when the display is going to start at the top of the screen and VIDEO which switches the electron beam ON and OFF. In practice, a fourth signal (intensity) selects whether the displayed dots are dim or bright. The display adapter generates these signals.

Colour monitor works just like monochrome one, except that they have three electron beams, one each for primary colours - red, green and blue. As the combination of primary colors produce others (red and green make yellow, blue and green make cyan and all the three colours mixed together make white), the entire spectrum can be created using three beams.

Electron beam in them cannot produce colour. The coating on a colour monitor consists of tiny dots of different phosphors that emit red, green or blue light when hit and energized by the electron beam. A finely perforated metal grid called a shadow mask is placed inside picture tube, just in front of the phosphor. It ensures that the electrons from each beam hit only the dots for which they are intended. Thus the red beam energizes only red dots and so on. In the following figure, we have shown the different switches of monitor that control its different activity.

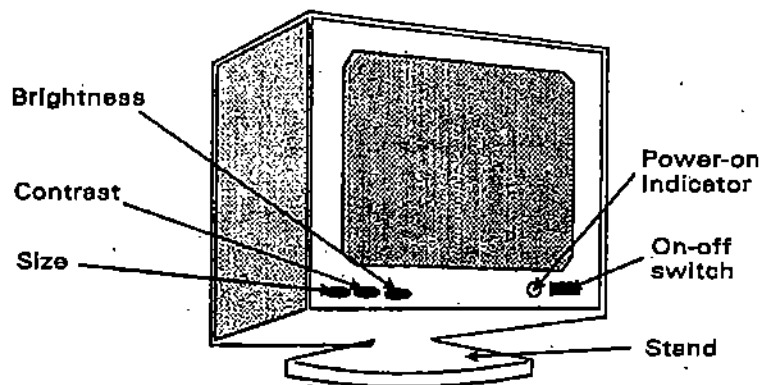


Figure of CRT Monitor

1.2.1 Pixels

This refers to an individual dot on the screen. All graphics you see on a PC are composed of hundreds (or even thousands) of these graphic dots or pixels. Each pixel can be of different colour thus.

1.2.2 Resolution

A resolution of monitor describes the number of potential pixels the monitor is capable of displaying. The value is given in horizontal and vertical pixels.

1.2.3 Raster Graphics

The form of video display system, which organizes the screen into a series of lines that is continually scanned number of times per second, is termed as raster display.

1.2.4 Interlacing

Information on a monitor is displayed by an electron gun scanning the phosphor screen inside of the display. The gun scans from top to bottom left to right with each complete scan displaying a frame. The gun works like a thin beam of spray paint, coating the inside of your monitor. In order to avoid a flickering image; some display adapters force the monitor to create an interlaced image. Instead of the electron gun scanning from top to bottom in a continuous manner, on the first pass it will skip every next line. On the second pass, it will scan the lines it missed during the first time, thus creating the full image in two scans instead of one.

If you are using resolutions of 800x600 or higher, you will need a monitor that displays information at these, resolutions in non-interlaced fashion to get a better image:

1.2.5 Non-interlaced Display

It is the opposite of interlaced display, For example all the lines are scanned or displayed in one pass instead of the two passes required in the interlaced display.

1.2.6 Frame Rate

This is used to show the number of times a screenful of information is produced per second or the number of times a frame is shown (in one second) on the monitor. Higher the frame rate, less flicker problem you get.

1.2.7 Video Bandwidth

It is the highest input frequency a monitor can handle and helps in determining the resolution capabilities of the monitor. It is measured in megahertz (MHz, one million cycles per second). If the bandwidth of input signal exceeds the given range of a monitor, it may cause compatibility problems and you may not get proper image 0:0 screens.

The higher the value, the better is the image; this frequency range is a factor determining the resolution of the monitor. The bandwidth does not represent the exact frequency at which the data is transferred instead it indicates the range in which the monitor can operate most optimally. If the input signal slightly exceeds the permitted range of the monitor, the visual effect is of softening of sharp edges of each pixel. In order to find out the minimum bandwidth required by a monitor to achieve a resolution, you must know the number of pixels and the frame rate related with that particular standard of display. A single signal pulse produces bandwidth the product of the total number of horizontal and vertical pixels and the frame rate assuming each pixel.

$$\text{Bandwidth} = \text{horizontal pixel} \times \text{vertical pixel} \times \text{frame rate}$$

For a 600 by 400 display if the frame rate is 25, bandwidth will be 5,760,000 or 5.7 MHz. actual bandwidth may slightly be higher than this (allowing for retrace time).

In the regarding of bandwidths, following standards are commonly used.

Video	Bandwidth (MHz)
MDA	16.3
CGA	14.3
EGA	16.3
VGA	25.2
8514/A	44.9

1.2.8 Synchronizing Signals

The electron beam in the monitor is spread across the screen by a combination of magnetic fields. One of the fields moves the electron beam horizontally while the other moves it vertically. Circuitry in the monitor supplies a steadily increasing voltage to

two sets of deflection coils to control the spreading of electron beam. These coils are electromagnets and the increasing voltage causes the field strength of the coils to increase and deflect the electron beam farther.

At the end of line, the field that controls the horizontal spreading of electron beam is instantly switched OFF, returning the beam to the starting side of the screen. Similarly, when the beam reaches the bottom of the screen, the field controlling the vertical spreading of electron beam is switched OFF, returning the beam to the top of the screen. The result is that the electron beam follows a tightly packed zigzag path from the top of the screen to the bottom and again to top to repeat this path.

1.2.9 Horizontal Scanning Frequency

The frequency at which the monitor repaints the horizontal lines that make up an image is called horizontal scanning frequency. Horizontal scanning frequency is measured in kHz and is standardized at 31.5 kHz for a VGA. For a Super VGA, this frequency ranges from 35 to 48 kHz, depending on the refresh rate of the video adapter.

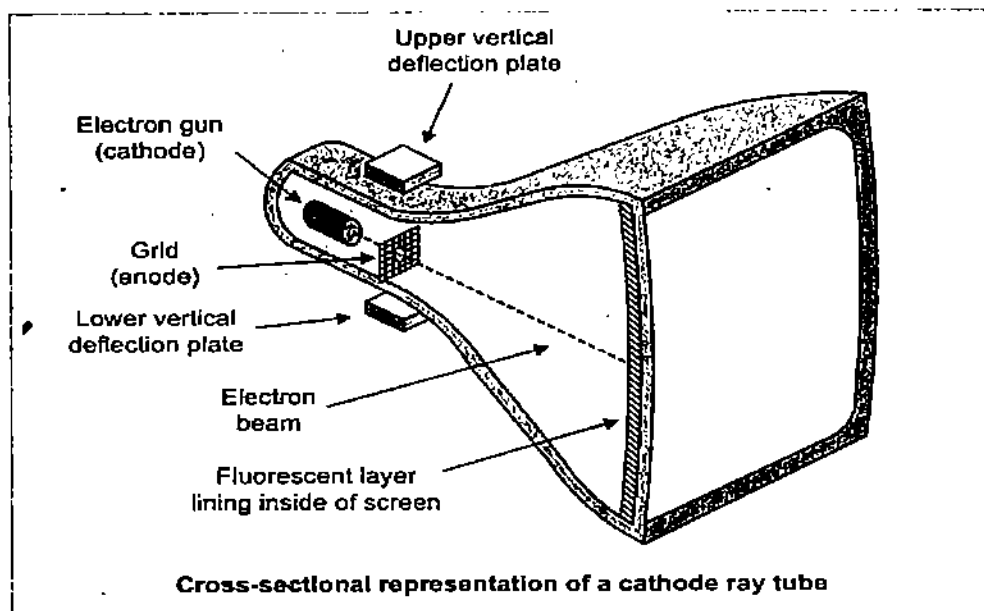
If your resolution is 800 x 600 i.e. if you have 600 lines per screen, and to avoid flicker if you show 30 frames per second (frame rate is 30). You will be showing 18,000 lines per second or you can say your horizontal scanning frequency is 18,000 Hz (Hertz) or 18 kHz (Kilo Hertz).

1.2.10 Vertical Scanning Frequency

The frequency at which the monitor repaints the whole screen, sometimes also known as the vertical refresh rate, is called the vertical scanning frequency. This is measured in Hz (numbers of cycles per second), and higher rates associate with less flicker. VGA has a vertical scanning frequency of 60 to 70Hz and Super VGA vary from the VESA guidelines of 56Hz (which is about the minimum tolerable) and 60Hz, to the official recommended standard of 72Hz (it means that screen should be redrawn 72 times per second).

The primary difference between the two steps is that several hundred horizontal lines are drawn for each vertical frame. Vertical frequency is also known as frame rate of the system because; one complete image frame is created every time the beam spreads from top to bottom. The electronics generating the spreading frequencies used by a monitor are inside the monitor itself. However, signals themselves must be synchronized with the data coming from the computer so that the characters appear at their proper positions on the screen.

In order to keep things organized, the video controller sends out special synchronizing signals, i.e., horizontal sync before each line is send to display and similarly vertical sync before each frame is send.



In the above figure, we have shown internal structure of CRT monitor.

1.3 OVERVIEW OF PICTURE TUBE

Picture on a picture tube is generated by continuous scan of electron beam. Inside the tube, Beam movement is just like our eye movement when we read a book, starting from the top left corner till the end of the line and then again back to the left corner but, to the next line and so on till the end of the page, then again start from the top of next page. In the same manner, picture tube generates the picture as follows:

The electron beam moves across the horizontal line generating all the picture elements of that line. At the end of the line, the beam returns to the left corner to next line or in other words the next vertical position. The returning line is called the horizontal retrace or fly back, and during this time the electron beam is switched off so that it should not disturb the screen image.

Using this horizontal and vertical scanning motion, the electron beam generates the image on the screen. Once a screenful of image is generated, at the end of frame, electron beam again moves to the top left corner of screen. This is called vertical retrace during this time also the electron beam is switched off so that it should not disturb the screen image.

1.3.1 Monochrome CRTS

The construction of a typical monochrome (one colour) CRT is shown in the following figure. Current flows through a heater, which makes the cathode hot, resulting in formation of a cloud of electrons around the cathode. The heated electron discharge from cathode is attracted by the accelerating anode. The focusing anode forces the electron moving towards the face of CRT to form a tight beam.

The accelerating anode increases the velocity of the electron beam so that it reaches the face of CRT in a very short time. This is necessary otherwise the electrons will repel each other rather quickly forcing the beam to spread and go out of focus. As the electron beam is moving towards the face of CRT, it passes through the deflection coils

and there are two pairs of deflection coils. One of the pair deflects the electron beam back and forth horizontally across the CRT while the other pair deflects it up and down vertically across the CRT.

The inner walls of the CRT are coated with phosphors, which emit visible light when struck by the high-velocity electron beam. The brightness of the images formed on the CRT face depends on the intensity of the electron beam when it strikes a particular spot. Applying a negative voltage at the grid of the tube controls the intensity of the beam. Presence of a high negative voltage at the grid prevents some of the electron from leaving the cathode (two negative forces repel each other) thus reducing the intensity of the beam and hence the brightness of the display.

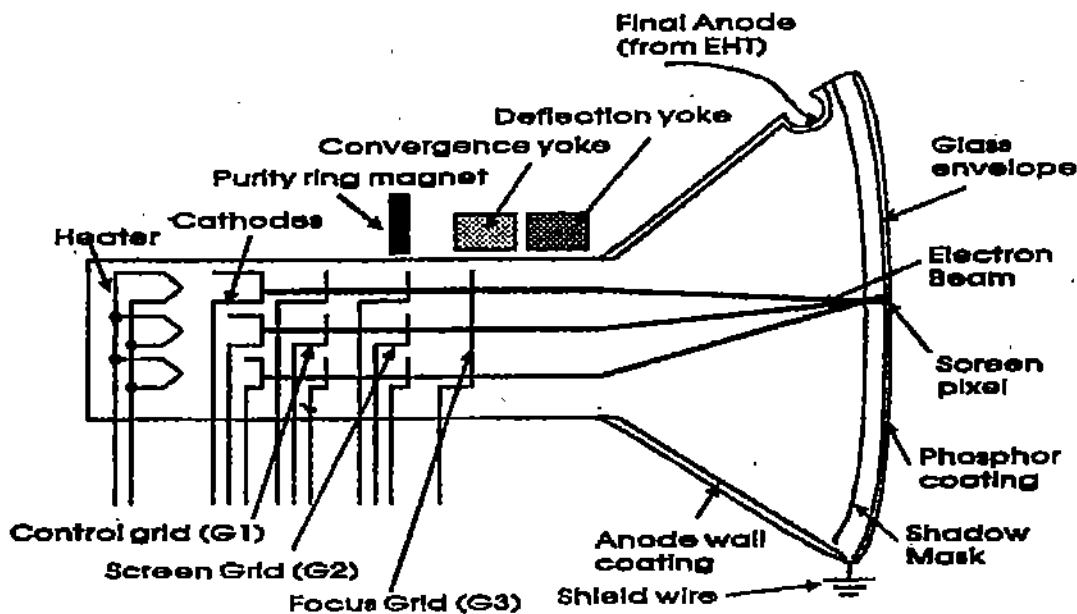
Applying a sufficiently high negative voltage at the grid will prevent any of the electrons from being emitted from the cathode and striking the phosphor coating. This is how the beam is turned off to produce a black portion of an image. The CRT sometimes use the principle known as electrostatic deflection, i.e., electrostatic field is produced by applying high voltages to horizontal and vertical control elements in the tube and are used to direct the electron beam to the desired position.

Electrons also have magnetic poles and may also be deflected by magnetic fields. In such cases, external coils of wires, which form the deflection yoke, replace the deflection plates. The deflection yoke is mounted near the neck of the CRT and it moves the electron beam to any desired location on the screen. Most of the computer monitor uses the magnetic deflection technique to produce a display. Since the basic principle of operation of colour CRT is similar with that of the monochrome CRT, they both are similar in many respects.

1.3.2 Color CRT

The main difference is that the colour CRT has three different colour phosphors on its CRT face, three electron guns and a shadow mask.

A simplified diagram of the colour CRT is shown in the following figure.



Deflection of the electron beam can either be electrostatic or electromagnetic.

The phosphor dots are arranged in such a manner that any group of three primary colour, i.e., red, green and blue, forms a dot. This group of three dots is also known as a triad varying the relative intensity of the three-phosphor spots in the group can produce different colours. You can produce any colour using different combination of red, green and blue. All the three different electron beams produced are deflected simultaneously by an equal amount and are not deflected independently. The focusing should also be such that the electron beams converge on the holes in the shadow mask.

When passed through a given hole, the beams deconverge and strike their associated phosphor dots, producing illumination of the three different colour phosphors, which make up a triad.

1.3.3. Resolution

Due to the nature of the raster scan process; all graphics and characters displayed will be comprised of various patterns of dots produced in a row at a time. These dots are known as pixel or pel. The horizontal resolution of a display depends on the horizontal frequency and the dot rate, which is determined by a dot clock. The dot rate is the rate with which the electron beams is turned on and off while moving across the screen. The number of dots that will constitute a line is determined by the dot clock in combination with the horizontal frequency. The number of dots that can be drawn on the screen can be roughly calculated by dividing the dot clock frequency by the horizontal frequency.

Vertical resolution is the number of lines displayed on the screen and it can be determined by dividing the horizontal frequency by the vertical frequency. However, if interlaced scanning is used, then a complete picture or frame requires two vertical oscillations as against the one required in the non-interlaced scan. Also, in case of interlaced scanning, any particular pixel is refreshed half as often as in case of non-interlaced scanning and hence higher persistence phosphors are required in the CRT, so as to avoid flickering of the display.

1.3.4. Shadow Mask

To produce a colour pixel just pointing the electron beam at the right dots is not enough because part of the beam will spill over and hit the other dots in the triad. The result of this spill over is a loss of colour purity. To prevent this effect and make images as sharp and colorful as possible, all colour CRTs use a shadow-mask (a metal sheet with fine perforations in it). It is located inside the display tube at a small distance behind the phosphor coating of the screen. The shadow mask and the phosphor coating on the CRT screen are necessary critically arranged so that through the shadow mask the red electron beam can hit red colour phosphor only. The other two colours of dots are in shadow of the mask and cannot be seen by the electron beam i.e. why we call it shadow mask.

The spacing of holes in the shadow mask, to a great degree determines the quality of the displayed image. For the spacing of the system to work properly, phosphor dots on the CRT screen must be spaced at the same distance as the holes in mask. Since the whole spacing determines the dot spacing, it is often termed the dot-pitch of the CRT; the term for slotted tubes is slot-pitch. (CRTs with in-line gun or Trinitron gun use slots rather than holes in the mask to achieve the same result).

The dot-pitch or slot-pitch of a CRT is simply a measurement of the distance between dots or slots. This measurement is independent of the size of the tube or the size of the displayed image. Closely looking at the screen of a colour monitor, you can see the shadow mask behind it. If you can see the holes in the mask, they are most likely .4mm or larger. On the other hand, if you cannot see the holes on the mask then they are likely .3mm or smaller.

1.3.5. Slot Mask (Aperature Grille)

Also known as aperature grille, it is a series of vertical wires held tightly in a frame to align the electron beams. They have the same function as shadow mask but achieve it in a better way. Sony Corporation introduced this technology. Aperature grill base monitor produce brighter and sharper image.

1.3.6. Screen Curvature

Most CRTs have a distinctive shape, a narrow neck that contains the electron gun or guns and fits inside the deflection yoke, a funnel like spreading that allows the electron to spread across the inner surface of the tube and the rectangular curved surface face. The common spherical curve makes sense for a couple of reasons. It makes the distance traveled by the electron beam more consistent at various points on the screen, edge to center and back to edge. A truly flat screen would require the beam to travel more at the edges than at the center and would require the electron beam to strike

1.3.7 True Resolution

The resolution of a video system refers to the details that it can display. It is a number of individual dots that make up the screen image and thus is a function of both the screen size and the dot-pitch. Since the size and number of dots, limit the image quality, the visible sharpness of screen images can be described by the number of dots can be displayed horizontally and vertically on the screen. For example, the resolution required by IBM's Colour Graphics Adapter is 640 dots horizontally and 200 dots vertically. The 8514/A Display Adapter produces an image 1024 by 768 dots in its highest resolution mode.

However, sometimes the resolution available on the screen made by a computer's display adapter are not equal. For example, a video mode of signed for the resolution abilities of a colour television set can hardly use the quality available from a computer monitor.

On the other hand, the computer-generated graphics may be designed for a monitor that's sharper than the one being used. For example, you might try to use a television in place of the more expensive monitor. The sharpness you actually see will be less than what the resolution of the video system is supposed to produce. Actual resolution is a physical quality of the video display system, the monitor, which is actually being used. It sets the ultimate upper limit on the display quality. In colour system, the limit on resolution is purely physical; it cannot be more than the dot-pitch of the tube. In monochrome systems, which have no shadow masks, the resolution is limited by the bandwidth of the monitor, the highest frequency signal with which it can deal. Finer details pack more information into the signals sent from computer system to the monitor. The more information in a given time, the more is the frequency of the signal.

1.4 Summary

In this unit, we have discussed about monitor, which is only responsible to produce output in the form of softcopy in every steps of operation of computer. We have also described different features of monitor. We have introduced its function of monitor. Finally, we have explained about picture tube.

1.5 Model Questions:

Write a short note on the following.

1. What are the roles of monitor? Can you explain its resolution?

2. How picture tube is important? Explain its vertical and horizontal scanning.

3. Describe the different feature of picture tube.

UNIT-2: DIFFERENT MONITORS AND THEIR VIDEO ADAPTERS

Structure:

2.0 Objectives

2.1 Introduction

2.2 Type of Monitor

2.2.1.1 Analog Monitors

2.2.1.2 Digital Monitors

2.2.2.1 Monochrome

2.2.2.2 TTL Monochrome

2.2.2.3 Composite Monochrome

2.2.2.4 VGA Monochrome

2.2.2.5 Multi scanning Monochrome

2.2.2.2 COLOUR MONITORS

2.2.2.2.1 Composite Color

2.2.2.2.2 RGB

2.2.2.2.3 Enhanced RGB

2.2.2.2.4 VGA Monitor

2.3 SPECIAL MONITOR (Laptop Display)

2.3.1 LCD Display

2.3.2 Gas-Plasma display

2.3.3 Dual Monitor Systems

2.3.4 Full page Display

2.4 VIDEO DISPLAY ADAPTERS

2.4.1 Basics of Video adapters

2.4.1.1 Video Cards

2.4.1.2 Standards and Compatibility with System

2.4.1.3 Suggested Memory for Board

2.4.2 Monochrome Display Adapter (MDA)

2.4.3 Colour Graphics Adapter (CGA)

2.4.4 Hercules Graphics Adapter (HGA)

2.4.5 Enhanced Graphics Adapter (EGA)

2.4.6 Professional Graphics Adapter (PGA)

2.4.7 Video Graphics Array (VGA)

2.4.8 Super VGA

2.4.9 8514/A Display Adapter

2.4.10. Extended Graphic Array (XGA)

2.5 Summary

2.6 Model Question

2.0 OBJECTIVES

At end of this unit, student will be able to:

- Understand about the basics of different type of digital monitors.
- Explain the different type of adapters.
- Distinguish between color monitors and LCD monitors.
- Know the concept of various feature of adapters used in the monitors.
- Analyse statistics of feature for suitability of adapters.
- Generate awareness about laptop display.

2.1 INTRODUCTION

The world of IBM-standard monitors is highly confusing. In order to be certain that you have the right type of display, you must decide it with clear-cut specification. Saying colour or monochrome is not enough. You must also indicate the signal standard to which the monitor must abide. The standard depends on the video adapter that is used by the computer, but some monitors work with different adapters and many adapters are flexible in regard to the monitor. Several factors determine whether a monitor will work with a particular display adapter. It must be able to synchronize with the adapter's horizontal and vertical scan rates.

The horizontal scan rate is the number of lines the electron beam can scan in a second. The higher the display resolution, the more scan lines are needed to form a complete picture.

The vertical scan rate (or refresh rate) is the number of complete screen per second. If it is too low, the eye can actually detect fresh as an irritating flicker, particularly when the picture has a light background.

At high resolutions, some adapters use a type of scanning called interlacing. Monitors that can sustain a high enough horizontal scan rate to display a picture of (for example, 1024x768 pixels 50 times a second), are expensive. If lower rate is used, it takes longer to display a complete picture and hence the vertical refresh rate also drops. This leads to unacceptable flicker. The hardware, compromises by scanning alternate lines on the first pass then going back and filling in the missing ones on the second pass. Although the complete picture is refreshed only perhaps 40 times a second, the illusion of an 80Hz refresh rate is created in greater way of reducing flicker. Another important characteristic of colour monitors is the dot pitch. It is the distance between individual phosphor dots and it effectively determines the smallest pixel that will appear sharply displayed on the screen.

For a typical high quality 14" VGA colour monitor, the dot pitch is 0.28mm. Cheaper ones may use larger dots with a pitch of 0.32mm.

2.2 TYPE OF MONITOR

2.2.1. According to their operating mode, following two type of monitor is popular in the computer field.

2.2.1.1 ANALOG MONITORS

This type of monitor is to be operated by analog signal. The advantages of using analog signals to control a monitor is that you can have an continuous variable range of colour because the signals that control the red, green and blue electron guns may themselves be continuous variable.

2.2.1.2 DIGITAL MONITORS

In digital monitors, the electrical signals from the adapter that control what appears on the screen are digital. In other words, they are either ON or OFF. With a digital monitor, you can combine the three primary colours each at only two levels of intensity - high (on) or low (off). This allows a maximum of 8 different colours, by adding another intensity signal we can increase this to 16 colours.

2.2.2. According to their looks, following two type of monitor is popular in the computer field.

2.2.2.1 MONOCHROME

Monochrome means exactly what it says, mono means one, and chrome indicates colour. Monochrome monitors show their images in one colour, be it green, amber, white, crimson. Among the monochrome monitors available you have three choices. A fourth, the multi scanning monochrome display, accepts almost any monochrome signal.

2.2.2.2 TTL Monochrome

The original display type offered by IBM - one that plugs into Monochrome: Display Adapter. It uses digital input signals and uses separate lines for its horizontal and vertical synchronizing signals and video signal. Its digital signals match the level used by integrated circuits of the Transistor-Transistor Logic family or TTL family. These chips operate with tightly defined voltage ranges indicating a logical one or zero. (Five volts is normally considered a digital one, although that is the input voltage level of TTL chips.

The maximum output level TTL signals ever reach is about 4.3 volts. Because of their use of TTL signals, such monitors are often called TTL monochrome displays. They can only be plugged into MDA (monochrome display adapter) or compatible displays adapters (including the Hercules Graphics Board).

2.2.2.3 Composite Monochrome

This type of monitor offers the lowest resolution of any monochrome system available for PCs, the same level as a CGA (red, blue and gray) display but without the colour.

Composite monochrome monitor uses the same signal as home and professional video systems. Designed for the mass market, the composite monochrome monitor is likely to be the least expensive available. It can be plugged into only a CGA or compatible display monochrome display today is in multimedia systems to preview video images.

2.2.2.4 VGA Monochrome

As like TTL monochrome monitors, VGA monochrome monitors follow IBM standard

A VGA monochrome monitor will work with any VGA display adapter without change. It will display VGA graphics without colour. As VGA monitor works on analog signal it can show unlimited number of gray shades whereas MDA and HGC were limited to black white and highlight.

2.2.2.5 Multi scanning Monochrome

Unlike other three monochrome display types, which are designed to operate at certain fixed frequencies or fixed standard, multi scanning monochrome display adapts to the signals sent to it within a wide range of frequencies.

Usually, this kind of monitor can handle any standard monochrome signal, from composite to VGA. It offers an advantage over fixed-frequency displays, that you can shift it between different monochrome-equipped computer systems no matter what standards they follow.

2.2.2.2 COLOUR MONITORS

Five types of colour display are generally available for connecting to PCs. Among these are composite colour, RGB (or CGA), Enhanced RGB (or EGA), VGA and multi- scanning monitors.

2.2.2.2.1 Composite Color

This is the kind of video monitor you are likely to connect to your VCR or video camera; they use the standard composite video signal. This signal standard has long been used with PCs - starting with the CGA adapter and PC's built-in display system.

Composite signals have never really gone away. They are still used where computer generated graphics are generated for television and video productions. They also link into multimedia systems. The 4.43MHz colour sub carrier specified by the PAL standard limits their colour sharpness, however, the best you can expect, should you want to use a composite colour display for general use, is readable 40-column text.

2.2.2.2.2 RGB

The original colour display for the IBM PC used three discrete digital signals for each of the three primary colours. From these three signals, the display type earned the name RGB from the list of additive primary colours: Red, Green and Blue. To be completely accurate, this type of monitor should be termed RGB; which is standing for intensity, as per CGA standard. Except for the interference signal, the RGB monitor works like a composite colour monitor, using the same frequencies. Because there is RGB monitors work with CGA, EGA (in its degraded CGA mode), and compatible display adapters as well as the PCs. Because of the low resolution of CGA systems, CGA monitors are dead and forgotten as the PCs computer monitor.

2.2.2.2.3 Enhanced RGB

Moving up to EGA quality requires a better display, one able to handle the 21.85 kHz horizontal synchronizing frequency of the EGA standard. In addition, its interface is somewhat different.

While still digital; it must accommodate intensity signals for each of the three primary colours. The EGA signals require a matching EGA connection on the display. So, EGA is essentially obsolete. No new systems are sold with it anymore. Rather than getting a new monitor to work with your existing EGA card is better to upgrade to VGA.

2.2.2.2.4 VGA Monitor

VGA display was introduced with PS/2 system. They use analog inputs and a 31.5 kHz horizontal synchronizing frequency, which is VGA standard. A VGA monitor with a VGA adapter can display from a palette of 2, 62,144 colours. At 320 x 200 resolutions it can show 256 colours and at a higher resolution of 640 x 480 it can show 16 colours. VGA's 640 x 480 mode is about 20 percent sharper than 720 x 348 mode produced by Hercules card. But the biggest advantage of VGA is it is compatible with CGA.

A digital monitor works like a digital clock, i.e. if you have only hour display on your clock it can show only hour values and seeing the display you can not guess in between values, if your display is 1, after one hour suddenly it becomes 2 and so on. You do not have any value in between values.

But, in your analog watch even if you have only one hand to show hour, as there is no sudden jump from 1 to 2, by seeing the position of the hand you can easily guess approximate minute value colors. In digital monitor, three colors are used RBG, each of these colours can have only two values. They are either ON or OFF and total combination you can have is as given below:

Red	Green	Blue
Off	Off	Off
Off	Off	On
Off	On	Off
Off	On	On
On	Off	Off
On	Off	On

On	On	Off
On	On	On

Each of this combination produces a single colour. As you can see you can have only 8 (2^3) different colours using three inputs. If you have four inputs you can have 2^4 or 16 colours and as you require more and more colour you require more and more wires. This becomes very complex to manage. This problem is not with an analog display like an analog watch analog display can have different in between values on its Red, Green and Blue wires, it is not restricted to ON and OFF value. For example you can have 0.5 volt or 0.55 volt or 0.555 volt on Red wire and each of these value will produce different Red colour VGA monitors are analog monitor so then can produce unlimited number of colours. VGA is now one of the computer monitor which is in demand.

2.3 SPECIAL PURPOSE MONITOR (Laptop Display)

CRT monitors are too big and heavy and consume too much current for laptop computers. Therefore, laptops use liquid crystal display (LCD) or glass plasma display that are light and draw little current.

2.3.1 LCD Display

Digital watch and calculator s uses LCD display. In it each pixel is an individual solid-state device, addressable by row and column co-ordinates.

In a LCD display each pixel is filled with a transparent material called liquid crystal, when energy is supplied, the liquid crystal material becomes polarized (i.e., it will pass/reflect light of a certain polarity only) and appear darker than its background or becomes opaque. In order to improve the contrast LCD's are usually back-lit or edge-lit. Still they are sometimes difficult to read at an angle or in bright sunlight.).

Polarized light passes through a liquid crystal sandwiched between transparent column and row electrodes. When current is passing through electrodes the molecules of the liquid crystal twist and they rotate incoming light by 90 degree. Next we have another polarizing filter, which is in 90 degree to the first filter. So, if light coming out of first filter is turned by 90 degree then only they can pass through second filter and show on display as a bright pixel.

If you do not apply any current to electrode, liquid crystal would not rotate incoming light and pass them as it is. The second filter will block that light and you get a dark or black pixel on display. In the following figure, we have shown the working of LCD monitor.

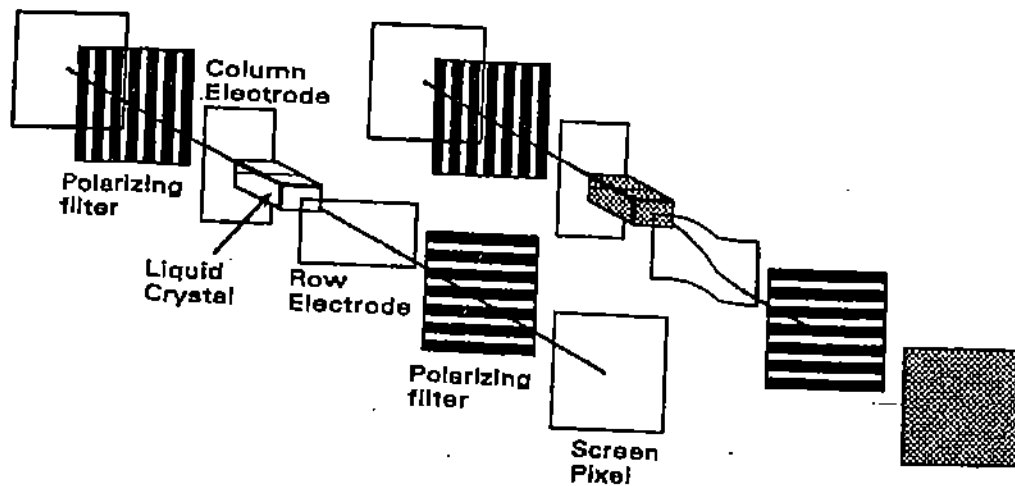


Figure: Steps of working of LCD Monitor basics

2.3.2 Gas-Plasma display

Gas-plasma displays produce a orange-red, brighter and sharper picture in comparison to the LCD's, but, require much more current and hence they are mainly used in portable computers which run on AC power being solid state, LCD and plasma display are reliable and rugged. Sometimes when a display driver circuit fails, making horizontal and vertical row of pixels either dark or light permanently. The only solution to this is replacing of the screen.

2.3.3 Dual Monitor Systems

Since, the monochrome and colour display adapters use different display memory areas and I/O port address, one can have two adapters and two monitors in use simultaneously. This would typically make sense in a computer aided design (CAD) system, where one might use a monochrome text display to enter commands and a. high-resolution colour monitor to show graphical output. In a dual monitor configuration, one display card must be an MDA, CGA or equivalent and the other may be anything. One adapter must be configured as the default display. This can be done through jumpers on the motherboard or setup program.

A VGA card must always be default display EGA or VGA cards, which can operate in either monochrome or colour mode, are restricted to modes that do not conflict with the alternate adapter.

2.3.4 Full page Display

Apart from moving towards higher resolution general purpose graphics standard, there have been special graphics standard aimed at particular applications. The most common is the A4 or A3 monochrome monitor designed for either DTP or CAD application. The resolution offered by A3/A4 is very variable because. The only real criterion that has to be met is that they display A3/A4 images. If you want, there is nothing stopping you from

displaying an A4 page using a 640x200 CGA display, only problem is you would not be able to see what is displayed on it. For example, some A4/A3 displays are based on Super VGA offering a resolution of only 800x600. These really do not deserve consideration for true A4/A3 display status. Super-VGA at 1024x768 offers a reasonable A4/A3 display but it can be unacceptably slow. A specially designed A4/A3 display usually involves a graphics, co-processor. This offers high resolution but at a reasonable speed. Resolutions for A4/A3 displays are anything from 1024x768 to 2048x1536 although 1024x1024 is more common.

The resolution used does not increase much in going from an A4 to a double sized A3 display. The key factor that changes is the size of the monitor. A large monitor (20" diagonal) can display a pair of readable A4 pages, even though the dot resolution is only 70 dots per inch i.e., about the same as a low quality dot-matrix printer.

In order to find out how a quality image you can expect, simply divide the number of dots across the screen by the screen width to give the dot per inch. Anything about 100 dots per inch range is good. There are other factors that affect quality. For example, image stability and the sharpness of each dot are important. Many displays also offer between 2 and 64 grey levels and this can, be used to provide smoother corners to the curves that make up a character. All of these factors make it very difficult to choose a monitor from a paper specification alone. If you are going to have to spend a long period of time working with an A4/A3 monitor then the best advice is to find a way of trying out a number of different models. As A4/A3 displays are usually standard they are often used with another monitor. A standard display adapter, usually VGA and monitor are use with an A3/A4 display. The operating system and supported application use the standard monitor and application for which drivers are supplied the big A4/A3 monitor.

Many users find having two monitors on their desk more than they can cope with. A better solution to this is to use an A4/A3 graphics card that supports at least one standard graphics mode. A4/A3 monitors generally come complete with drivers for the best selling DTP and CAD packages. The packages that are most often supported are - Ventura, PageMaker and AutoCAD. It is also usual for Windows and GEM drivers to be included. One important point is to make sure that the drivers supplied actually work with the latest version of the programs that you are going to use. For example, when Ventura became available in Version 3 Professional edition it took some time for the A4/A3 makers to produce working drivers.

2.4 DISPLAY ADAPTER

During the early days of IBM PC and compatibles development cycle, choosing the video system was an extremely simple one. The Monochrome Display for those who did not want colour and the Colour Graphics Display was available for those who wanted colour. Since then, a number of adapters and display options have come into existence, giving you a wide range of choices. In the area of video adapters and displays, it is very important to follow the industry standards. Many video systems are not supported by every program and system peripheral.

Of all your computer's components, the monitor is the most visible. You can not see how the 4MB of RAM or a 260MB hard disk drive looks like from outside, but a large colorful monitor is bound to impress anyone. In a way, the monitor can be termed as the mouth of a PC: It displays information, enabling you to interact with the computer and to know what is actually taking place at present. Although it looks to be a single unit, you must

know that there are actually two components that make up the entire computer display system. They are the monitor and the display adapter.

The monitor is only half of the equation of video in the PC. The other, more important, half is the display adapter (also known as video card), the electronic circuitry that controls the monitor. It is actually an expansion card plugged into the motherboard inside the main unit.

The display adapter enables information to leave your Computer and appear on the monitor. The adapter cards contain the special circuitry that drives the monitor, telling it what to display, where to display it and what colours to use etc. Since the majority of PC's do not have built-in video adapter; the display adapter is an option. Ideally, you should choose one based on your text, video and graphics need.

Note: that there are special monitors for each display adapters. You cannot (and should never) mix and match them carelessly.

In some PC's, especially the laptops, the display adapter is built into the computer. The display adapter may also be called the display adapter card, video card, video hardware or the video system.

Text and Graphic Mode Display

There is a division in adapters between graphics and text modes. Adapters running in a text mode see the screen as a grid of rows and columns, with each intersection consisting of a cell that is able to hold a single character or graphic block shape. Each cell can be made up of 49 to 126 or more pixels or dots. The computer and its programs need only to specify which character will be in each cell. The adapter circuit will then look for that character's pattern of pixels in a special ROM chip (Character generator ROM). The final result is a series of signals sent to the display to form letters, numerals, punctuation or tiny graphic shapes. Text mode does not use much memory because it only sees the screen as a set of approximately 2000 (80 columns by 25 rows) cells. Each cell requires only a single byte to record and store its character. However, more bytes will be required for colour and other attribute information about that character.

A typical screen can be held in 4K of memory, allowing enough room for color and other attribute (blink, underline, reverse etc.) information. Adapters running in a graphics mode see the screen as grid of pixels. Graphics image and text can be building from the individual pixels.

Text made this way is known as bit-mapped text, because it is built entirely from the individual bits and pixels instead of from pre-formed character shapes inserted into cells. Both graphics and text constructions are stored in memory before being sent to the display. More pixels in the field mean higher resolution. The latest resolutions are from 640x480 to 1024x768. As each pixel needs at least a single bit of memory (for monochrome), or as many as three bytes of memory (for millions of colours), graphics modes use much more memory than the text modes do.

For example a 640x480 monochrome needs nearly 40K. An 800x600 display of 256 colours needs 512K. That is not the memory of 512K in your PC that is just for the graphics other than any other memory in your computer. Often the number of colours trades off against the resolution. The highest resolution mode would not offer as many colours as a lower resolution mode because there is not enough memory to show both the top resolution and the top colour possibilities at the same time.

2.4.1 Basics Video adapters

2.4.1.1 Video Cards

A computer and a monitor without any video display card attached to it, is of no use for us. A display adapter or a video card is must for transferring data that is to be displayed on the screen from the processor to monitor.

Video add on cards are available in three basic groups, the 8-bit, the 16-bit and the 32-bit. These terms refer to the amount of data that can be transferred at one single pass. A 16-bit card can send data to the monitor in half the time required by an 8-bit card. It is important to have the exact form of video card required by your particular computer. The older PC's and XT's usually have an 8-bit card, the newer AT system require a 16-bit card. It is the most common form of card at present.

2.4.1.2 Standards and Compatibility with System

There are applications where a monochrome monitor and the display card will be fine but there are also applications requiring high resolution for sharp colour display, and for that you must have an expensive colour monitor and appropriate display card. Various video display standards are available in the market. Some of them are discussed below. Although the monitor displays the picture, it is the adapter which actually generates it. Adapter is a card that plugs into the expansion bus, though on many new machines it is built on the motherboard.

Several display standards are in wide use. All adapters are essentially similar in that they are memory-mapped, i.e., they reserve an area of memory to hold a representation of the screen. Besides memory, display adapters have hardware I/O ports. Software writes to them to do things such as clear the screen or set the position of cursor and size etc. Software can also read information from them. For example, it can learn the current size of the cursor or whether the CRT electron beam is doing a retrace. Various display adapters usually require different displays because they use different horizontal and vertical scanning frequencies.

Although some newer monitors like multi scan or multi sync, support multiple scanning frequencies, for a display to work with a particular adapter, the display must match the desired scanning frequencies otherwise an in-appropriate adapter can destroy the monitor.

2.4.1.3 Suggested Memory for Board

Most VGA boards nowadays can display 800x600 modes. The better and quality VGA boards will support 1024x768, either in interlaced or non-interlaced mode. More resolution means more dots on the screen, which means that the video board needs more memory.

Two things determine a video board's memory requirements: its resolution and the number of colours that it can display. For colours, but in the higher 640x480 resolution, it can only display 16 colours.

This has nothing to do with the monitor or the VGA board, except for the amount of memory on board. A normal VGA comes with 256K of memory right on the board. Resolutions like 1024x768 with 256 colours obviously require more memory.

Following is the amount of memory (RAM) that a video board needs for the common resolution/ colour combinations.

Mode	Resolution	Memory
MDA	80X25	14K
CGA	640 x 200	16K
	160 x 200	
HGA	720 x 348	64K
EGA	640 x 350 with 16 colours	64K to 256K
VGA	640 x 480 with 16 colours or 320 x 200 with 256 colours	256K
Super VGA	640 x 480. with 256 colours or 800 x 600 with 16 colours or 1024 x 768 with 4 colours	256K to 1MB
8514/A	11024 x 768 with 16 colours	512K to 1MB
XGA	1280 x 1024 with 256 colours	1MB

2.4.2 MONOCHROME DISPLAY ADAPTER (MDA)

The original and simplest display combination is IBM Monochrome Display Adapter and IBM Monochrome Display Monitor. Introduced in 1981 by IBM, this adapter could display only text (no graphics) in one colour, at a resolution of 640 pixels width by 350 pixels height. Because of its clarity and high-resolution, 640x350 pixels, it is best suited for word processing work. With 9x14 dot character box, monochrome monitor displays clear and attractive characters.

Pin Configuration of Monochrome Display Adapter Cable

Pin Number	Signal Direction	Description
1		Ground
2		Ground
3		Not Connected
4		Not Connected
5		Not Connected

6	>	+ Intensity
7	>	+ Video
8	>	+ Horizontal
9	>	-Vertical

Dot Box

For legibility IBM set the character box for the MDA at 9x14 pixels with a typical character using a 7x9 matrix in the box. The extra dots are used to keep individual line separate for greater readability. To put this character box on the screen in a default arrangement used by most VDTs 80 columns and 25 rows, requires 720 pixels horizontally and 350 vertically, a total of 252,000 dots on every screen.

Frame Rate

IBM compromised in how to display all these dots. At a high frame rate, displaying all 252,000 dots would require a wide bandwidth monitor than was available when the PC was introduced. IBM slowed down the frame rate to 50Hz and compensated for any flicker that might develop by using long persistence phosphors in its standard monochrome display. The lower frame rate gave the horizontal sweep of the scanning electron beam time to cover each line of the image. However, even with the lower frame rate, the dot density of the IBM monochrome standard demanded a higher horizontal frequency (18.43 kHz) than was used by popular video monitors (15.52 kHz).

MDA Summary

Resolution	Colours	Mode	Character	Vertical	Horizontal
72Q x 350	4	Text	80 x 25	50Hz	18.43KHz

2.4.3 COLOUR GRAPHICS ADAPTER (CGA)

IBM introduced this video adapter in 1981, which provided low-resolution text and graphics. CGA provided several different text and graphics modes, including 40 or 80 columns by 25 line, 16-colour text mode, and graphics modes of 640 horizontal pixels by 200 vertical pixels with 2 colour, or 320 horizontal pixels by 200 vertical pixels with 4 colours.

This adapter has basically two modes of operation; they are Alphanumeric (AIN, Text mode) or All Points Addressable (APA, Graphic mode). In AIN mode, the card operates in a 40 column by 25 lines or in 80 columns by 25 lines mode with 16 colours. In both these modes, character set is generated with a resolution of 8x8 pixels. On the other hand, in APA, mode two resolutions are available; they are a four-colour medium resolution mode of 320x200 pixels and a two-colour high-resolution mode of 640x200 pixels.

With Colour Graphics Adapter, you can choose from a number of monitors because the horizontal scanning rate of colour graphic board is the industry standard or 15.75 kHz. The

distance between the dots, making up the display is known as Dot pitch, the smaller the pitch, the sharper is the image. The IBM PC colour monitor has a pitch of 0.43mm.

Most of the colour monitor, available with IBM PC is RGB and not the composite monitors. The colour signal input to a composite monitor contains a mixture of colours that needs to be decoded and separated by the monitor. RGB monitors receive red, green and blue signals separately and combine these colours in required proportions to create other colours. RGB monitors offer better resolution than the composite monitors.

Pin Connection of CGA Composite Video Cable

Resolution	Colors	Mode	Character	Vertical (Hz)	Horizontal (KHz)
320 x 200	16	Text	40 x 25	60	15.75
640 x 200	16	Text	80 x 25	60	15.75
160 x 200	16	Graph		60	15.75
320 x 200	4	Graph	40 x 25	60	15.75
640 x 200	2	Graph	40x25	60	15.75

2.4.4 HERCULES GRAPHICS CARD (HGC)

CGA display systems had two major drawbacks. It was not particularly sharp and it required the purchase of both a video adapter and a new monitor, to add graphics to an original PC. The latter need was particularly troublesome if you had just purchased a monochrome system or had no need for color.

Adding a second display for graphics when system already had a perfectly good and sharp, text-based system appeared unproductive. The Hercules Graphics Card or HGC was the obvious solution, adding bit-mapped graphics to the character-mapped MDA.

Compatibility

The foundation of the HGC was completely copying of the MDA functions. Two boards worked exactly alike in the text mode, operated at the same frequencies with exactly the same cabling and display. Characters were formed in the same 9x14 pixel dot-box on a screen with a full-screen resolution of 720x348 pixels a horizontal a frequency of 8.43 kHz and a 50Hz frame rate. All attributes of the IBM MDA underline, blink, high-intensity and reverse video are supported by the HGC. HGC even included a parallel printer port, same as the MDA.

Although the HGC is compatible with MDA hardware and its text-mode software, it is not compatible with any IBM graphics standard. Applications must be specifically written with support for the HGC. Programs written for the CGA and other IBM graphics standards will not execute properly on the HGC unless they have HGC support.

Memory Requirement

The difference between MDA and HGC was its memory. Instead of four kilobyte of the MDA, the HGC was equipped with a full 64K, arranged in two 32K banks. That is,

occupying the ranges assigned to both MDA and the CGA boards. The various video modes supported by the HGC allowed the use of this memory for several purposes. In text mode, it could be used for number of text pages up to 16. In graphics mode, it was sufficient for two full-screen pages of high-resolution monochrome image.

Hercules monochrome system's 64K was too large to fit into 32K space reserved for monochrome memory. It conflicted with any CGA board plugged into the same system.

As a solution Hercules provided software HGC.COM with the board. In default, boot-up mode, HGC would activate only half of its memory (32K), eliminating the conflict. Running this program with its FULL option put the full memory of the HGC into use. The half option has switched off the upper half. Later software included this procedure in software driver designed for Hercules graphics.

Graphics Display Standard

In its graphics mode, HGC provide a very good resolution of 720x348. One bit is assigned to each pixel, allowing no attributes (other than the pixel being OFF or ON). However, lines are not stored in memory in the same order that they are displayed on the screen. Contiguous lines in memory display four lines apart on the screen. In effect, the screen is divided into four fields, and one line from each field in sequence is fitted into memory, then the second line from each field, and so on. Switching to graphics mode is accomplished by software configuration of a switch register. The first bit of this register controls the mode. As you have seen biggest drawback of HGC is it cannot show any attribute, you get plain black and white image.

2.4.5 ENHANCED GRAPHICS ADAPTER (EGA)

The enhanced graphics adapter (EGA) was the card, which started the progress towards higher resolution colour display. Before EGA, the display standard available was either monochrome or the 4-colour CGA (colour graphics adapter). EGA has a resolution of 640x350 with 16 colours. EGA became standard PC graphics adapter around 1986; a couple of years after it were introduced.

IBM originally offered EGA cards with PC/AT, but as CGA cards were cheaper and more widely supported, it took time for EGA. Software to appear and by then EGA standard had been replaced by The Enhanced Display System consists of a graphics board, a graphics memory expansion board, a graphics memory module kit and a high resolution colour monitor. A positive thing about EGA card is that it can be used with IBM monochrome display; IBM colour display or the new IBM enhanced colour display.

EGA, an IBM colour monitor can display 16 colours in either the 320x200 or the 640x200 modes. With EGA IBM monochrome monitor, give resolution of 640x350 pixels with a 8x14 character (text mode) box. Through software, the character box can be manipulated up to the size of 8x32 and the 16 colours can be selected from a palette of 64.

The 256-character set can be enlarged to 512-character set by using the IBM-memory expansion card. A 1024 character set is added with the IBM graphics memory module kit. These characters sets are loaded from programs. The enhanced graphics' colour adapter has a maximum of 128K memories that fits into the RAM space just above the 640K limits.

In case you have installed more than 640K of memory; you probably will be losing the extra memory after the installation of EGA.

The graphic's memory expansion card adds 64K of memory to the standard 64K thus making the total 128K. The IBM graphics memory module kit adds another 128K for a total of 256K. This second 128K of memory is only on the card and does not consume any of the PC's memory.

EGA Summary

Resolution	Colours	Mode	Character	Vertical (Hz)	Horizontal (KHz)
320 x 350	16	Text	40x25	60	21.85
640 x 350	16	Text	80 x 25	60	21.85
720 x 350	4	Text	80 x 25	50	18.43
320 x 200	16	Graph	40 x 25	60	15.75
640 x 200	16	Graph	80x25	60	15.75
640 x 350	4	Graph	80x25	50	18.43
640 x 350	16	Graph	80 x 25	60	21.85

2.4.6 PROFESSIONAL GRAPICS ADAPTER (PGA)

IBM introduced the Professionals Graphics Display System in 1984. The system is composed of a Professional Graphics Monitor and a Professional Graphics Card Set. When fully installed, this card set uses three slots in an XT or AT system. The PGA offers three-dimensional rotation, and clipping as a built-in hardware function, the adapter can run 60 frames per second of animation, because the Professional Graphics Adapter uses a built-in dedicated, microcomputer. The PGA has an 8088 microprocessor, 320K of RAM and 64K of ROM. The resolution of this system is 640 by 480 pixels. The PGA system was designed for application such as Computer Aided Design (CAD). The professional Graphics card and monitor were targeted towards the engineering and scientific applications rather than towards financial or business applications.

This system was discontinued when the PS/2 was introduced and has been placed by the VGA and 8514 graphics standards for these newer systems.

PGA Summary

Resolution	Colours	Mode	Character	Vertical (Hz)	Horizontal (KHz)
320 x 200	16	Text	40 x 25	60	15.75
640 x 200	16	Text	80 x 25	60	15.75
320'x 200	4	Graph	40 x 25	60	15.75
640 x 200	2	Graph	80x25	60	15.75
640 x 480	256	Graph		60	30.48

2.4.7 VIDEO GRAPHICS ARRAY (VGA)

VGA stands for video graphics array and it uses a special video chip to generate display. IBM, the producer of first such chip, uses a VLSI (very large scale integration) chip comprising of a huge number of logic gate arrays to make video circuitry. The outcome is an image in graphics mode with resolution of 640x480 pixels and a display of 16 colours from a palette of 2, 62, 144. It was a high-colour and high-resolution graphics adapter built into IBM's PS/2 computers, introduced in 1987. With the introduction of VGA, one could still run applications written for earlier display standards; however, you could not use the same monitor. The only monitor capable of handling the vertical and horizontal frequencies set by the new VGA standard is the multi scanning monitor or VGA monitor. VGA offers more graphics, more colours and highest resolution of any graphics adapter before it. It has a lot of features and flexibility. In fact, the VGA adapter should be the choice for anyone considering a colour adapter for his or her system.

As far as graphics mode is concerned, VGA is very versatile. It is fully compatible with MDA, CGA, and EGA etc. VGA cards are widely duplicated by other manufacturers and are cheap. The clones offer features that even IBM's VGA lacks. Some VGA clones even offer Hercules graphics emulation and can be used with a monochrome monitor. VGA has its own monochrome, or "shade of gray", mode that looks beautiful on some of the paper-white monitors.

The VGA BIOS (Basic Input Output System) is the control software present in the system ROM for controlling the VGA circuitry. With the BIOS present, software can give command and function to BIOS without requiring manipulating the VGA directly. The VGA can run almost any software that has been originally written for MDA, CGA or EGA. If you use a monochrome monitor with VGA, Colour summing to 64 gray shades is done in ROM BIOS summing routine is initialized if the BIOS detects the monochrome display when the system is booted. This routine uses a formula that takes the desired colour and rewrites the grey intensity value. The colour that would be displayed, for example, is converted into 30% red plus 59% green plus 11% blue to achieve the desired grey.

The VGA uses up to 64 shades of grey when converting colour modes of operations for display on the IBM Monochrome Display. Users who prefer a monochrome display, therefore, can execute colour-based applications.

VGA Summary

Resolution	Colours	Mode	Character	Vertical (Hz)	Horizontal (KHz)
360 x 400	16	Text	40 x 25	70	31.5
720 x 400	16	Text	80 x 25	70	31.5
320 x 200	4	Graph	40 x 25	70	31.5
640 x 200	2	Graph	80 x 25	70	31.5
720 x 400	16	Text	80 x 25	70	31.5
320 x 200	16	Graph	40 x 25	70	31.5
640 x 200	16	Graph	80 x 25	70	31.5
640 x 350	4	Graph	80 x 25	70	31.5
640 x 350	16	Graph	80 x 25	70	31.5
640 x 480	2	Graph	80 x 25	60	31.5
640 x 480	16	Graph	80 x 25	60	31.5
320 x 200	256	Graph	40 x 25	70	31.5

2.4.8 SUPER VGA

Super VGA refers to any enhancements to existing IBM VGA standard. The original standard set by VESA was that of 800x600 resolution and 16 colours. However, today cards are produced with 256 colours and even 1024x768 resolutions. However, the problem lies in the number of graphics display modes, the VGA standards itself have 17 different ways to display an image on the monitor. The numbers of colours on higher resolution increase in the newer graphics display standard like the Super VGA, IBM's 8514/A adapter and the extended graphics standard (XGA).

The new generation video adapters using the continuous edge graphics (CEG) chips can even deliver 7,40,000 colours simultaneously on the monitor. More specialized standards can deliver a resolution as high as 1600x1200 and are able to generate millions of colours. However, the first thing to be considered before buying a display adapter is the software, i.e., whether the software carries the necessary driver so as to make the software function properly with respect to the display card you have in your computer system.

Although most of the software now come with the VGA standards, along with the downward compatibility, i.e., compatible with older important graphics display standard like the CGA, EGA, MDA, etc., be sure that the software you purchase have the required driver. Another factor to be kept in mind is that whether the monitor you have is capable to cater the high resolution being offered by the display card. The older display standard compatible monitors used the digital interface, but the VGA needs an analog interface.

Only the multi scanning monitors have the facility to offer both the digital as well as the analog interface but it is not necessary that the entire multi-scanning monitor will work with the VGA display adapter. The monitor you are having needs to be fully capable to function with the specified horizontal and vertical frequencies specified by the particular graphics display adapter.

As VGA became the accepted new standard, board manufacturers their attention first concentrated on building compatible boards and they turn their attention on expanding the capabilities of basic VGA. Any VGA board that offers modes beyond the basic IBM VGA is

non-Standard. Such board and operating modes are generally referred as Super VGA (1988). If you opt for a Super VGA board then you are risking very little as you can always use it as a standard VGA board.

Software Drivers for Super VGA

However, it is important to realize that Super VGA modes offered may not be usable unless you have the correct type of monitor and software that supports the board. Your super VGA acts like a super VGA only when you have got drivers program to support super VGA. Efforts are in process to establish standards for Super VGA. The best known is VESA VGA BIOS extension.

This defines three new modes.

800x600 in 16 colours

640x480 in 256 colours

800x600 in 256 colours

At present VESA (Video Electronics Standards Association) A group of VGA vendors) has little impact, but most Super VGA boards support these modes. A non-VESA graphics mode that is commonly found in Super VGA boards is 1024x768 in up to 16 colours but VESA has plan to include this in the standard.

You can find other non-standard modes that give even higher resolution but notice that, making use of these may be more troublesome than it worth. Especially so when you remember that a VGA card, even a Super VGA card is a dumb graphics device and all the work, drawing, etc has to be done by the processor. Even with fast 386/486 CPU this can involve moving very large quantities of data within the video memory.

In other words, there comes a point where increasing the resolution of a VGA fails because of the burden this places on the processor. Most people consider that 1024x768 is about as far as VGA should go.

VGA and TV (PAL)

Few years back, a composite video monitor was the standard display screen for home computers. Composite monitors have not been used much with PCs. Now, many devices are available which allows the display of computer images on TV monitors or which can capture pictures from video camera. However, these high-resolution devices are supported by very limited number of programs and are very costly as well.

The question is however, why do such need arise?

Firstly, large composite Video monitors and rear projection screens cost far less than an analog display screen of same size.

Secondly, the output can be sent to VCR and hence allowing you to make tapes of desktop presentation graphics or animation film or even one can record the screen output of an entire seminar.

One of the limitations of the VGA-TV cards is that they do not have GENLOCK capability to synchronize the Video with the external source. You need GENLOCK to be able to overlay image from the live VGA-TV output. External devices are required to here are programs available which makes it possible to capture screen from television, videotape running in a VCR or from a video camera. Once you capture an image you can save it as gray-scale TIFF, EPS or Publisher's Paintbrush etc. format. So that, these

images can be used in desktop publishing or video database and also once, the image is captured, it can be manipulated and adjusted in many ways. There are controls for varying brightness and contrast and one can even create special effects with the image manipulation software.

2.4.9 8514/A DISPLAY ADAPTER

In 1987, IBM introduced the 8514/A display card and the monitor offering a 640x480 resolution with 256 colours and an interlaced 1024 x768 resolution with 16 or 256 colours. 8514 stands for the monitor number and the /A is for the display card or adapter. However, IBM itself is now trying to supersede this video standard with its latest extended graphics array (XGA) display cards. The PS/2 Display Adapter 8514/A offers higher resolution and more colours than the standard VGA.

This adapter is designed to use the PS/2 Colour Display 8514 and plugs into a Micro Channel slot in any PS/2 model. An operation and modes of the built-in VGA continue to be available. An IBM Personal System 8514 Memory Expansion Kit is available for the IBM Display Adapter 8514/A. This kit gives increased colour and grey scale support. Offering 7,86,432 pixels at 1024x768 resolution, extended VGA more than doubles a screen's information potential in comparison with standard VGA's 3,07,200 pixels at 640x480 resolution. Almost two-third of the high-performance VGA cards available is able to display this impressively high resolution.

Although this resolution is same as that of the more costly 8514/A graphics adapter, it does not mean that they compatible with 8154/A. This is so because 8514/A are designed to work with channel bus and also use an in-built graphics microprocessor.

Besides this, the extended VGA and 8514/A creates this resolution in different manners. The display generated by the 8514/A is interlaced, i.e., scanning across alternate lines of the screen and hence requiring two pass to produce the image generate the image. Standard 8514/A and extended VGA boards outfitted with 512K RAM memory can give the same number of colours in the 1024X768 resolution mode, however; 8514/A with added 512K RAM can display 256 colours, something definitely beyond the capability of VGA cards. Although VGA boards with added 512K RAM can display 256 colures only at a resolution mode of 800x600.

The main difference between extended VGA boards and 8514/A is the graphics co-processor with 8514/A. With the co-processor with it, 8514/A boards can move the graphics processing from the main microprocessor to co-processor, thus achieving faster performance as co-processor, freeing the main microprocessor, does the creation of graphics.

Apart from boosting the graphics performance, it also frees the main microprocessor almost immediately once the command is passed to the graphics co-processor. This off-loading of main microprocessor can substantially enhance type performance of the system.

8514/A Summary

Resolution	Colours	Mode	Character	Vertical (Hz)	Horizontal (KHz)
1024 x 768	256	Graph	85 x 38	43.4	35.52
640 x 480	256	Graph	80 x 34	60	31.5
1024 x 768	256	Graph	146 x 51	43.4	35.52

2.4.10 Extended Graphic Array (XGA)

IBM's latest video display standard is the extended graphics array (XGA), and it contains both the VGA and the 8514 level of resolution on the same board. However, it is interlaced and does not support the non-interlaced mode. Latest XGN2 announced in September 1992 can work in non-interlaced mode and offers more number of colours at almost twice the speed of XGA hi. 1990 IBM announced the Extended Graphics Array (XGA). It is the display adapter for 80486-based systems in PS/2 family and offers high performance through the use of graphic co-processor.

XGA has three modes of operation, viz., standard VGA mode, an enhanced VGA 132-column text mode and its own XGA mode. In XGA mode, the display resolutions are equivalent to those provided by the 8514/A, but offers 256 colours at 1024x768 and up to 65,536 colours at 640x480 resolution. However, XGA is incompatible with previous PC display modes at hardware level, hence it requires special drivers. The 65,536 colour, 640x480-resolution mode offered by XGA shows almost photograph like output. For this resolution 16 bit pixels are used. 5 bit for red (32 shades of red), 6 bit for green (64 shades of green as eye is more sensitive to green) and 5 bit for blue (32 shades of blue).

VIDEO SPEED

Display cards that we use in the PC are almost all dumb; the CPU must do all the work. For instance, if a program wants to put a picture of a circle on the screen, CPU must compute the location of every dot on the screen, and activate the corresponding pixel, one by one. Most advanced system uses an object-oriented approach. The CPU just issues a command to the (intelligent) video board: video board draws a circle. Place it here, and colour it blue. The CPU could describe an entire screen with a few commands to the video board and go back to computing while the graphic hardware handles the display work. Video speeds a real problem in the PC performance, especially in graphic based environment like Window, what can be done about improving video performance?

There are several approaches:

- Faster video drivers.
- A video board with a co-processor or an accelerator.
- A local-bus video board that will exploit the full speed available from your computer.

Faster video displays, i.e., putting the information on the screen as soon as possible, are soon for just about any PC. But, it is critical for graphic application such as animation, desktop publishing, drafting and presentation etc. Speed or video performance becomes more important with higher resolution and more colours on screen. A colour Super-VGA screen can, means moving about one fourth MB of information. There are several memory factors to video speed like 8-bit vs. 16-bit, VRAM vs. DRAM and ROM BIOS vs. BIOS shadowing.

2.5 Summary

In this unit, we have discussed about different type of monitors. We have also described different type of video adapter used enhance picture quality of monitors in computer field. We have introduced the installation of printer using Windows' Add Printer Wizard. Finally, we have explained about LCD and Gas-plasma display used as monitor.

2.6 Model Question:

Write a short note on the following.

1. What are the roles of software monitors? Can you explain different type of digital monitors?

2. What do you understand by color monitors? Explain how it is better as compared to monochrome monitors.

3. Describe the different type of video adapters.

UNIT-3: INTERFACING AND UPGRADING THE MONITOR

Structure:

- 3.0 Objectives
- 3.1 Introduction
- 3.2 Upgradation of Video Adapter
 - 3.2.1 Removal of Cover of PC
 - 3.2.2 Installation of a Video Adapter
 - 3.2.3 Replacement of Cover of PC
- 3.3 Upgradation a Monitor
- 3.5 Monitor Handling
 - 3.5.1 Cleaning the Monitor
 - 3.5.2 Prevention Phosphor Burn
- 3.6 Summary
- 3.7 Model Question

3.0 OBJECTIVES

At end of this unit, you will be able to:

- Understand about the basics of upgrading of the monitor.
- Explain the proper interfacing of the monitor.
- Know the concept of installing the video adapter.
- Analyse statistics of prevention and handling monitor.
- Generate awareness about phosphor burn.

3.1 INTRODUCTION

Most of the computer user starts with a monochrome monitor and Hercules display adapter. But, as the requirement change and as the cost of higher end displays are coming down, you may want to upgrade your display system.

We find that a combination of dual-frequency monochrome monitor and dual graphic adapter (combination of HGA and CGA) is the most cost-effective combination with the best possible display. Next option if you can afford is VGA or multi scanning mono monitor With a VGA card and if you can afford you must go for SVGA or Multi scanning colour monitor with SVGA card.

3.2 Upgradation of Video Adapter

The method of installing a new card or replacing an old display card is as given below.

1. Remove cover of PC.
2. Remove old card if you are upgrading.
3. Set dip switch, jumper etc of new card.
4. Press down the new card in slot.
5. Setup Computer and all the software you have for new card.
6. Close the case.

3.2.1 Removal of Cover of PC

In order to remove the lid (cover) of the console, do the following steps.

- Switch off the power being given to the computer and turn off all peripherals as well.
- Remove the screws on the back of the case of PC and place them in a dish or other suitable place.
- Remove power supply cable both from the rear end of console as well as from the wall socket.
- Remove the lid, which slides forward. This should be done in a very smooth manner. Take enough care not to press the disk drives when pushing the lid forward and neither should the lid get entangled with any cables system present inside the computer. Once the lid is forward enough, lift it up and pull off.

3.2.2 Installation of a Video Adapter

- Find out the old video adapter connected in the motherboard. It is easy to locate the video card as the monitor cable is connected to it.
- Remove the cable connecting the monitor with the video card. If the printer cable is also attached, remove it as well.
- Take off the screws connecting the video card with its mounting slot.
- Take out the video card and place it in a safe place.
- Before connecting the video card-make sure, it is configured properly. Set switches on the new card and set any jumpers or DIP switches and if required add RAM on it as well.
- Place the card on the mounting slot. In case you are upgrading your present video card from an 8-bit one with a 16-bit card, you need to install a new mounting slot. Finally, check all the connection.
- Attach the cable between the monitor and the card.

On a XT or higher system set DIP switch on the motherboard according to the display you have installed.

Card	DIP switch		Monitor
	Switch 4	Switch 5	
MDA/HGC	Off	Off	Monochrome
CGA	Off	On	CGA 80 Column
CGA	On	Off	CGA 40 Column
EGA	On	On	No Monitor
VGA	On	On	No Monitor
SVGA	On	On	No Monitor

You must set EGA, VGA as No monitor because they have their own BIOS. In case you have an AT system, you will need to run the SETUP program if there is any error on switching on the system.

3.2.3 Replacement of Cover of PC

- Properly align the lid (cover) over the case.
- Smoothly push the cover back and again take enough care no to entangle the cover with the cables present inside the case otherwise. They may get cut or defected.
- Tighten the screws back on the lid.
- Reassemble the PC.

3.3 Upgradation of Monitor

Connecting a monitor to the system is a fairly simple task. The thing to be taken care of is that the monitor and the display adapter must match. The steps normally followed to add a monitor to the system are as described below. If nothing is seen on the monitor, check whether power to the system is switch off. One should never plug or unplug anything with a computer power supply on. Though it is possible to plug a monitor into a powered system, it is safer to do the connection in the off state. Remove the monitor cable attached with the display card and place the old monitor in a safe place. Place the new monitor on top of the main unit or any other place you would like.

Connect the monitor cable with video card. Plug in the monitor and switch on the computer. In case the monitor and the display adapter do not match, a POST error will be displayed. This error may be due to a faulty DIP switch setting in case of PC/XT systems. In case of an AT system, run the SETUP program to rectify the fault. The monitor is properly plugged in and is switched on, as most of the monitor has a power switch of his or her own. Also make sure that the brightness on the screen is proper.

If you hear some squeaking sound or smell or see smoke switch off the monitor immediately and consult a proper service person.

3.5 Monitor Handling

3.5.1 Cleaning the Monitor

Monitors attract a lot of dust, and you may find fingerprints and other mess on the screen. To clean your monitor, dip a towel into some soap water. Squeeze water from the towel or the water may drip inside monitor and short-circuit its circuits. Next, switch off your monitor and computer and now softly clean monitor surface.

3.5.2 Prevention Phosphor Burn

Leaving your monitor on all the time leads to something called phosphor burn-in and after sometime, the same image becomes "etched" on the screen. You may see Lotus1-2-3 or WordStar even with the monitor turned off. This happens because displaying the same image for a period of time burns up the phosphor lining inside your display. When you are not using the PC turn-off, your monitor avoids phosphor burn-in. Even if you will be away from the PC for a few minutes, it is a good idea to turn down the brightness. This saves the monitor, keeping the image as brilliant as it was when new.

Another way to avoid phosphor burn-in is to use a screen dimmer or screen blanker software. These special programs blank or erase the screen after a given amount of time, usually when you have not touched the keyboard or mouse for a predefined time. Pressing any key on the keyboard or moving mouse restores the image.

Windows contains its own screen saver. You can find it in the Control Panel, the Desktop item. Other screen blanking are also available, the most popular among these being the after Dark application. It shows lot of special effects on your screen.

3.6 Summary

In this unit, we have discussed about upgradation of video adapter, which is more important for quality improvement of picture. We have also described installation of video adapter. Finally, we have introduced proper handling of monitor, which are necessary for keeping life long.

3.7 Model Question:

Write a short note on the following.

1. What is meat of video adapter? Can you explain its upgradation in brief?

2. Write about various steps of the upgradation of Monitor.

3. How to you will handle the monitor? Describe in brief.

UNIT- 4: MAINTENANCE OF MONITOR AND ITS TROUBLE-SHOOTING

Structure:

4.0 Objectives

4.1 Introduction

4.2 Screen Problem

4.2.1 Monitor dead

4.2.2 Unsteady display, moving band on screen

4.2.3 Display out of sync

4.2.4 Display rolling in horizontal direction

4.2.5 Display rolling in vertical direction

4.2.6 Fold-over in the beginning or end of scans line

4.2.7 Brightness control

4.2.8 Spot seen when switching off

4.2.9 Black vertical bar on the screen

4.2.10 Defocusing of dots

4.2.11 Weak contrast

4.2.12 Raster ok and No video

4.2.13 Noisy display

4.2.14 Fine noise on the screen (white spot)

4.2.15 Corrupted display

4.2.16 Distorted characters on screen

4.2.17 No cursor is Seen

4.2.18 CMOS Display Type Mismatch

4.2.19 Corners display are improper

4.2.20 Screen display is displaced to one side

4.2.21 Tile display towards one direction

4.3 Summary

4.4 Model Question

4.0 OBJECTIVES

At end of this unit, you will be able to:

- Understand the basics of screen problem, which generally arise in the monitor.
- Explain the general troubleshooting of the monitor.
- Know the concept of various types of troubleshooting of display mismatching in the monitor.
- Analyze statistics of noise prevention in the monitor.
- Generate awareness about weak contrast of monitor.

4.1 INTRODUCTION

Today, in the field of computer, assembling and maintenance are the most important problems. The programmers who are developing software must have deep knowledge of hardware for programming well. Now, here we discuss some clues about troubleshooting and maintenance of monitor.

4.2 Screen Problem

Whenever we catch a screen problem, the first step is to find out what could be the cause of the problem. Secondly, we search that either it is due to incompatible of display card or monitor, if it is incorrect configuration or it is due to faulty monitor or faulty display adapter card.

The display card and monitor are quite reliable but a gradual appearance of faults with age is very common. Many display related problems are caused by incorrect configuration and cable fault. Another common reason of blank display could be monitor's power switch being OFF or the brightness being turned down.

There are many things that can cause display problems. Some of the most common faults are listed below:

4.2.1 Monitor dead

- Fuse of monitor has blown.
- Power section of monitor has been faulty.
- Fault in horizontal section.

4.2.2 Unsteady display, moving band on screen

- Interference in supply line.
- Power supply voltage very low.
- Power supply circuit faulty.

4.2.3 Display out of synchronization

- Cable not connected properly.
- Vertical/horizontal hold misadjusted.
- Monitor synchronization circuit's needs adjustment. . Display mode not supported by monitor.
- Wrong software driver for display adapter.

4.2.4 Display rolling in horizontal direction

- Horizontal hold control misadjusted.
- Loose cable contact.

4.2.5 Display rolling in vertical direction

- Vertical hold not adjusts properly.
- Loose cable contact.

4.2.6 Fold-over in the beginning or end of scans line

- Phase control preset misadjusted.
- Fault in vertical output or horizontal output section (component value deviated).

4.2.7 Brightness control not working, abnormal or no change in brightness

- Brightness control defective.
- A normal grid voltage (G2).

4.2.8 Spot seen when switching off

- CRT becomes very weak.
- CRT becomes very old.

4.2.9 Black vertical bar on the screen

- Control grid voltage is not proper.
- Control grid voltage circuit is defective

4.2.10 Defocusing of dots

- Improper focus voltage.
- CRT deterioration.

4.2.11 Weak contrast

- Fault in video circuit

- Fault in contrast control circuit.

4.2.12 Raster ok and No video

- Video cable contact not proper.
- Fault in video circuit.

4.2.13 Noisy display

- Problem in grounding.
- Loose video cable contact.

4.2.14 Fine noise on the screen (white spot)

- It causes problem in connectors.
- Leakage in high voltage protection.

4.2.15 Corrupted display

- Software corrupts display adapter memory.
- Display adapter memory (RAM) faulty.

4.2.16 Distorted characters on screen

- Fault in Display adapter.

4.2.17 No cursor has Seen

- Software fault
- Adapter not fully IBM-compatible.

4.2.18 Display Type Mismatch

- CMOS setup incorrect for type of display, run SETUP.

4.2.19 Corners of the display are improper:

- Adjust ring magnets located on deflection yoke.

4.2.20 Display on the screen is displaced to one side:

- Centering magnet of deflection yoke misadjusted

4.2.21 Display tilted towards one direction:

- Deflection yoke tilted due to loose fitting.

4.3 Summary:

In this unit, we have discussed about trouble shootings of dot matrix printer, which is more popular for its very low cost consumables which are too used. We have also described different trouble shooting of Ink jet printer, which is also popular for its quality. Finally, we have introduced trouble shootings of Laser printer, which is popular for its cost and quality.

4.4 Model Questions:

Write a short note on the following.

1. Can you explain trouble shootings of Monitor?

2. Write down about the factor of weak contrast of monitor.

3. Describe the effect of noise that disturbs the display.



Uttar Pradesh
Rajarshi Tandon Open University

BLISS-03

**Computer Network and
Security Maintenance**

BLOCK

8

Maintenance and Trouble Shooting Of Printers

Unit 1

Overviews of Different Printers and their Working **5**

Unit 2

Interfacing and Installing the Printers **18**

Unit 3

Servicing and Trouble-Shooting the Printers **31**

Unit 4

Working with Plotters **42**

BLOCK INTRODUCTION

Now printer and plotter is such common device by which we can capable to take hardcopy of our output. Day by day, some of advancement in the computer field is going to improve the technology. However, as we know the latest type of output devices like printer and plotter also going to improve their technology. Therefore, we have needed to study such devices in the field of computer.

Now in this block, we will discuss about previously mentioned output devices. We will also explain about installation and configuration of these devices. Finally, we consider the description of maintenance and trouble shooting of these devices that gives the real boost to generating hardware engineer those must become aware about printing and plotting devices used in computer field..

We will mainly focus on following point:

Concept of Installation of Printer and Plotter: Concept of installation printer and plotter are more necessary, when candidates are wanted to became maintenance engineer in computer field must be aware about previously mentioned devices.

Concept of Configuration Printer and Plotter: Concept of configuration of printer and Plotter are also necessary, when candidates are wanted to became maintenance engineer in computer field will require knowing fact of configuration of Printer and Plotter.

Concept of Maintenance Printer and Plotter: Concept of maintenance Printer and Plotter are necessary, when candidates are wanted to hold the knowledge to become computer maintenance engineer.

Concept of trouble shooting of Printer and Plotter: Concept of trouble shooting of Printer and Plotter is compulsory for candidates who are wanted to do hardware maintenance in computer field.

Unit wise description of block is given as below:

The first unit of this block is concentrated on basics of different type printer, which create conceptual view of function of different printer.

In the second unit of this block, we have devoted on interfacing of printers. We further discussed about its maintenance.

In the third unit of this block, we have mainly discuss about servicing and trouble shooting of printer, which are also essential for computer professional.

In the fourth unit of this block, we have described about installation and maintenance of plotter.

Suggestion for further reading

1. PC Upgrade & Maintenance Guide, Fifteen Edition, Mark Minasi
2. Modern All about Printer, Manhar Lotia, Pradeep Nair
3. Inside the PC, Eight Edition, Peter Norton

UNIT1: OVERVIEWS OF DIFFERENT PRINTERS AND THEIR WORKING

Structure:

- 1.0 Objectives
- 1.1 Introduction
- 1.2 Type of Printers
 - 1.2.1 Impact Printers
 - 1.2.1.1 Dot Matrix Printer
 - 1.2.1.2 Daisy Wheel Printer
 - 1.2.1.3 Line Printer
 - 1.2.1.3.1 Drum Line Printer
 - 1.2.1.3.2 Chain Printer
 - 1.2.1 Non-Impact Printer
 - 1.2.2.1 Inkjet Printer
 - 1.2.2.2 Laser printer
 - 1.2.2.3 Thermal Printer
 - 1.2.2.3.1 Color Thermal Printer
 - 1.2.2.3.2 Thermal-Wax Transfer
 - 1.2.2.3.3 Thermal Dye-Diffusion
- 1.3 Working Process
 - 1.3.1 Printing
 - 1.3.2 Paper Handling
 - 1.3.2.1 Friction Feed
 - 1.3.2.2 Automatic Feed
 - 1.3.2.3 Rolling Feed
 - 1.3.2.4 Pin-feed and Tractor-feed
 - 1.3.2.5 Push and Pull Tractors
 - 1.3.3 Paper Control
- 1.4 Summary
- 1.5 Model Questions

1.0 OBJECTIVES

At end of this unit, student will be able to:

- Understand about the basics of different type of printers.
- Explain the various kinds of printers and its components.
- Distinguish in between impact printers and non-impact type of printers.
- Know the concept various types of interfacing devices used to connect printer.
- Analyze statistical feature for suitability of printer accordingly our purpose.
- Generate awareness about working process of different printers.
- Identify the various clue of paper handling.

1.1 INTRODUCTION

A printer is an output device used to operate under control of computer. Most of computers uses printer to produce permanent output (say hardcopy) in human readable form. Printer is produce printed reports and document such as bank statements, invoices, pay slips as well as hard copy of some graphics

1.2 Type of Printers

1.2.1 Impact Printers

1.2.2 Non Impact printers

1.2.1 impact Printers

The basic principles of these printers are having physical component containing the character. These printers are having standard set of character, sizes and standard font. In this characters are hitting against the ribbon and embossed on paper. Following printers are to be considered as impact printer.

1.2.1.1 Dot Matrix Printer

It is an impact type of printer. It prints one character at time, characters are building them by dots, each dot being formed by firing the solenoid wires passing through holes and embossed on paper. The size of character matrix is of two types 5/7 and 7/9. Quality of output is poor because characters are formed by dots. The typical speed of the matrix printer ranges from 180 characters per second (cps) to 360 cps. These printers are available in 80 col 132 col or 136 col. In the following figure, we have shown the printer.



This is one of the most common printers in the PC environment; number of Dot-Matrix printer's user may exceed the users of all other printers put together. Some other printers such as Laser printer may provide better result, or the line printers may give

faster output, but a Dot-Matrix printer is an all-purpose, low-maintenance printer with facility to print text and graphic on almost any type of paper.

Most Dot-Matrix printer models are inexpensive, small, easy to operate and very efficient at their job. As the market demand continues, printer manufacturers introduce newer, faster and less noisy Dot-Matrix printers so; there is every chance that Dot-Matrix printers will stay with us for the years to come.

Let us see how a Dot-Matrix printer prints. The main image formation part of a Dot-Matrix printer is the "print head". On the print head of Dot Matrix printer there is a number of small print-pin mechanisms.

A print-pin mechanism is a combination of permanent magnet, electromagnet, spring and print pins. These print pins are typically arranged as shown in the given figure. Print head is available in 9 pin, 18 pin and 24 pin models.

- 9-pin models have one column of nine pins,
- 18-pin models have two columns of nine pins each, while
- 24-pin models have three columns of eight pins each, although some machines use different arrangements.

Most of the low cost personal computer printers use 9-pin print head, but to produce high quality output, an increasing number of newer impacts Dot-Matrix printers are using 18 or 24-pins. When a line of character is to be printed, the print head moves horizontally across the paper and print-pins strikes the paper through an inked ribbon.

The pins impact is precisely timed so that it strikes on the right position in the character matrix at the right time. A major factor in determining the printing speed of a Dot-Matrix machine is the time required between successive strikes of each print-pin. The time needed to retract and reactivate each print-pin puts a physical limit on how rapidly the pins can fire. This limits the speed with which the print head can move across the paper. Because the head cannot go to the next dot position before each of the print-pins are retracted and are ready to fire. To speed up the printing operation, most of the impact Dot-Matrix printers print bidirectional, i.e. printing one row from left to right and then the next row from right to left. This saves the time that would ordinarily be wasted to move the carriage to the left side of the page to start the next line. Printers with this bi-directional feature must have sufficient memory to store a full line of text so that it can be read backwards and printed. Most of the impact Dot Matrix printers generate the character image based on the bit-image information stored inside its ROM memory.

Let us see step-by-step printing process of a Dot-Matrix printer:

- First, the PC sends a series of ASCII codes that represent characters, punctuation marks etc to be printed to the printer over the serial or parallel cable.
- With this PC also sends some printer movements information such as tabs, carriage returns and form feeds etc., which controls the position of the print head and print carriage.
- The ASCII codes received by the printer are first stored inside a buffer, which is a small random access memory (RAM) inside the printer. This buffer memory is used because it usually takes longer for a Dot-Matrix printer to print characters than it takes for a PC to send those characters

to the printer, the buffer helps free up the PC so that it can perform other functions during printing.

- The internal buffer of a Dot-Matrix printer generally has only 8 to 10 Kilo Byte (KB) storage capacity. When the buffer gets full, the printer informs the PC to stop sending any further character to the printer until some of the characters in the buffer is printed.
- Once the buffer frees up some space by printing the characters, the printer informs PC to resumes sending data. The processor inside the printer selects a particular dot pattern for the ASCII code received from the PC, from the bitmap table contained in the ROM (read-only memory) chips of printer.
- This table tells the printer the pattern of dots it should use to produce the character shape represented by the received ASCII code. The processor of printer takes the dot pattern information from the bitmap table and sends the signals to the print head; based on this information the print head fires different pins on the print head.
- The printer processor also controls the movements of the print head and the plate to move paper. The print head contains 9, 18 or 24 wires, called print-pins that are aligned vertically.
- One end of each of the pins is connected to individual solenoid, or electromagnet.
- To fire the print wire against the ribbon and paper, this electromagnet is energized (the current from the processor activates this electromagnet), which creates a magnetic field that repels a magnet on the end of the pin, causing the pin to move towards the paper.
- The moving pin strikes a ribbon that is coated with ink. The force of the impact transfers ink to the paper on the other side of the ribbon.
- After the pin fires, the electromagnet is de-energized and the spring pulls the print-pin back into its original position, ready to fire again.
- The print head is mounted on a track so it can move back and forth across the paper surface. The print head continues firing different combinations of print wires as it moves across the page so that various vertical dot patterns can be printed to makeup different character shapes.

Print quality of Dot matrix printer can be improved by the print head making a second pass over the same line to print a second set of dots that are slightly offset from the first set.

This magnet and spring printing mechanism is designed with one main purpose, i.e. to hold the print-pins away from the paper when no power is supplied to the printer and the print head, otherwise the print-pins may stray out of the print head and may get damaged.

This complexity of the print head can be justified by the protection it provides to the delicate print-pins.

1.2.1.2 Daisy Wheel Printer

It is also impact printer type character printer. In this printer plastic daisy wheel is used, which travels across the width of the printer. The quality of the output is good. This is similar to electronic typewriter. It has standard size of fonts; set of characters,

relatively daisy wheel printer speed is slow than dot matrix printer. The maximum speed is 180 cps to 280 cps.

1.2.1.3 Line Printer

Line printer is also impact type of printer. Such printer prints one line at a time. There are two types of line printer.

- Drum / Barrel printer
- Chain printer

These printers are heavy-duty printer used in large-scale organization.

1.2.1.3.1 Drum Line Printer

In this printer character is mounted on the surface of the drum. When drum rotates clockwise and anti clockwise direction, each character is repeated along the length of the drum. The width of the printer is 160 to 254 characters in a line, during the printing. Drum is revolves at high speed. For each character position there is hammer and between the hammer and characters a wide carbon ribbon and line print paper the paper is in a continuous sheet, which can be separated into individual pages after printing. In this up to 1 + 7 carbon copies printouts can be taken. The speed of such printer is ranges from 900 to 1500 lines per minuets.

1.2.1.3.2 Chain Printer

In the chain type of printer, the characters are mounted on the outside of a revolving chain, which moves fast in either direction. Relatively drum printer and the chain printers are slow in speed. The typical speed range from 600 to 900 lines per minutes and mechanism is most similar to drum printer.

1.2.2 Non-Impact Printers

1.2.2.1 Inkjet Printer

Inkjet printers occupy a position between Dot-Matrix printers and the Laser Printers. These are small printers that provide the resolution of the Laser printers at a cost just close to the Dot-Matrix printer. But, per page printing cost of the Inkjet printer is much higher than the Dot-Matrix printers; it is even higher than that of the Laser printers.

These printers are perfect compromise of cost, speed and quality. Only with the close inspection one can detect the difference between Laser printer output and the Inkjet output. Another quality of the Inkjet is that these printers print very silently; one does not hear the noise associated with the Dot-Matrix printers.

For their similarity to Laser printers, the Inkjet printers really work like a Dot-Matrix printer. They both have print heads that travel across the width of the page, printing an entire line of text with each pass. This mechanical movement puts Inkjet printers in the same category as the impact Dot Matrix printers.

As the Inkjet printers print character by character, the speed of the Inkjet printers is specified in CPS (Characters per Second), common speed is 100-200 CPS.

The biggest difference between Inkjet printer and the Dot-Matrix printer is the Inkjet's print head. An Inkjet printer spits little drops of ink onto paper to produce an image. It is a technology that works much better than one can imagine, let us see this complete printing process in detail

- An ink-filled print cartridge attached to the Inkjet's print head moves sideways across the width of a sheet of paper. This sheet of paper is fed through the printer below the print head.
- The print head is made up of small ink-filled chambers, each attached to a nozzle smaller than a human hair.
- An electrical pulse flows through a heating element i.e. a thin resistor located at the back of the ink chambers.
- When an electrical current flows through the resistor the resistor heats a thin layer of ink at the bottom of the chamber to more than 900° F for a very short duration to several millionths of a second. This makes the ink to boil and forms a small bubble of vapor.
- As this vapor bubble expands, it pushes ink through the nozzle to form a droplet at the tip of the nozzle.
- When the bubble further expands, the ink droplet overcomes the surface tension of the ink and the pressure of the bubble forces the droplet onto the paper.
- A typical character is formed by a dot matrix of these drops.
- As the resistor cools, the bubble collapses and the resulting suction pulls fresh ink from the attached reservoir into the ink chamber.

For color printing multiple cartridges of three basic colors Cyan, Magenta and Yellow are used. The color output of the Inkjet printers are very high quality because there is no ribbon to fade the color, ink is directly deposited on the paper. Also, by mixing different colors on the paper Inkjet printers can provide millions of colors. Most of the good quality Inkjet printers provide one additional black cartridge other than three basic color cartridges. This black cartridge is used to print black text or other black color images. Mixing of the three basic colors to produce black gives very faded black color.

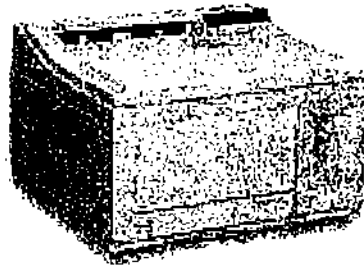
Some Inkjet printers use piezo electric crystal in place of the resistor to force the ink out of the chamber. When the electric current is applied to these crystals, it switches, forcing the ink out of the chamber.

1.2.2.2 Laser Printer

In a Laser printer, the Photosensitive Drum or the Organic Photo Conductive (OPC) drum is the main component for the image formation system. It is the center of the complete image formation process. This OPC drum is contained in a Toner Cartridge, which also contains toner, developing area, drum cleaning area and the primary Corona. Laser printer contains parts that deteriorate with use, as most of these parts are inside the Toner Cartridge, by replacing the Toner Cartridge, these parts are

replaced. When someone refills an empty cartridge, instead of replacing it, the print quality deteriorates.

In the following figure, we have shown the laser printer.



The OPC drum has special properties, which allows an image to be formed on its surface and then transferred to paper or some other media. This drum is an aluminum cylinder, which is coated with a nontoxic layer of Organic Photo Conductive (OPC) material, which gives it the name of OPC Drum. Aluminum base of this drum is electrically connected to the ground potential. The OPC material has properties similar to a photodiode, when it is exposed to light it becomes electrically conductive (in one direction only, like a diode). When exposed to light, any charge on the surface of the drum, conduct to the aluminum base of the drum (connected to ground i.e. to zero potential). Areas of the drum not exposed to light remain nonconductive and retain the charge.

As the OPC drum is sensitive to the light, it should never be exposed to direct sunlight or any other bright light source; as this could permanently damage the drum.

Let us now see the steps performed by the Laser Printer during the print process.

- **Charge:** This step cleans the surface of the cylindrical photosensitive OPC drum and places a uniform negative electrostatic charge on it.
- **Expose:** A data modulated laser beam scans the drum surface, converting selected negative points to positive points.
- **Develop:** Dry toner powder is attached to the positive points of the drum.
- **Transfer:** Image made with dry toner powder on the drum surface is transferred to the piece of paper.
- **Detach:** This step removes the piece of paper, along with the dry toner image on the paper surface, from the drum.
- **Fusing:** These steps fix the toner image on the paper with heat and pressure.
- **Clean:** Cleans any residual toner from the surface of the drum and prepares the drum for the next print cycle.

1.2.2.3 Thermal Printer

Thermal printers are another non-impact bit-image type of printers. The image generation method of these printers are same as that of the Dot-Matrix printer, only

difference is that instead of the pins striking ribbon, here pin is heated and brought in contact with special paper.

Thermal printers use a resistance that heats up with the flow of current. In the case of the thermal printer the resistance element is very tiny which heats and cools quickly, in a fraction of a second. Thermal printers actually burn the paper they print on. That would be dangerous. Instead, thermal printers use special, thermal sensitive paper that turns from white to near black at a moderately high temperature.

Thermal technology is ideal for portable printers because there are a few moving parts. Only the print head move. There are no moving parts inside the print head. There are no springs and wires to jam. Moreover, the tiny resistive elements require very little power to heat. Thermal printers can be lightweight, quiet and reliable. They can even run on batteries.

The special paper required for thermal printer has one drawback. Not only it is costly but also it is prone to discolor if it is inadvertently heated to too high a temperature.

The paper cannot tell the difference between a hot print head and heat from any other source. With time, image on this paper fades, to keep a permanent copy one need to take Xerox copy of its printout. Low-cost Inkjets are replacing the thermal printers from its position and these printers are slowly becoming available only on special-purpose machines.

1.2.2.3.1 Color Thermal Printer

Presently, two thermal printing method thermal-wax-transfer and thermal-dye-sublimation are most common thermal color printing technology.

As these technologies are almost same except the printing ink transfer sheet used in them, some hybrid thermal printers are also available which can be used as both dye-sublimation printer and wax-transfer printer by changing the ink transfer sheet.

1.2.2.3.2 Thermal-Wax Transfer

Thermal-Wax printers are used to produce good quality color printouts.

Current thermal-wax printers can print in resolutions as high as 300 DPI and produce colors equal to that of the photographic image.

These printers use wide ink transfer sheets, pure colors supported in a wax-based medium clinging to a plastic film base. The ink is transferred from these transfer sheets to the paper by a thermal print head as wide as the paper on which the image is being printed.

A series of thermal elements, 300 of them per inch, heat the transfer sheet, melting the wax to the paper and allowing it to resolidify and firmly fix to the paper.

Compared to other technologies, thermal-wax, printers are slow and wasteful.

- They are slow because the thermal printing elements must have a chance to cool off before advancing the 1/300th of an inch to the next line on the paper.
- They are wasteful because a transfer sheet as large as the sheet of paper is used for each of the primary colors printed on each page that is nearly four feet of transfer sheet for each page printed.

This makes printing a full-color page very expensive. Some thermal-wax printers give the user choice of three or four-pass printing. A three-pass transfer sheet holds the three composite colors of ink i.e. cyan, magenta and yellow, while a four-color sheet has one more color i.e. the black. It although black color can be made by overlaying the three primary colors, a separate black ink gives richer, deeper black tones.

It also imposes a higher cost and extends printing time by one-third.

1.2.2.3.3 Thermal Dye-Diffusion

For true photograph quality output from a printer, the thermal dye-diffusion process is considered better than the thermal-wax-transfer process. Using a mechanism similar to that of the thermal-wax-transfer process, dye-diffusion printers are designed to use penetrating dyes rather than wax. Instead of a dot merely being present or absent as in the case of a thermal-wax printer, diffusion allows the depth of the color of each dot to vary.

Dye-diffusion printing use ink transfer sheet that contain translucent colored dyes instead of wax. Printing head vaporizes the dye, which then condenses in the coating of the special paper used for printing. On this type of printer by varying the temperature of the print head precise amount of the dye for each dot can be controlled, this allows the variation in each color to be reproduced by changing the amount of dye required to print the image, rather than by changing the dot size as done by most of the other printing method. This allows the dye-diffusion printing to provide almost photograph like color output.

One drawback of this printer is the limited size of the output. Most of these printers provide only a postcard size output i.e. the output of most dye-diffusion printer's looks like photographs in size as well as color. Another limit is cost, print out with these printers are expensive.

1.3 Working Process

1.3.1 Printing

The color thermal printer feeds a sheet of specially coated paper into the print engine, where a roller that presses the paper against a wide ink transfer sheet coated with colored inks mixed with wax or special dye holds the paper.

The ink transfer sheet contains a band of each of the composite printing colors cyan, magenta, yellow and black. Each color band covers the width and length of the sheet of paper being printed.

Thermal wax printing use an ink transfer sheet covered with thin layer of colored wax, the print head melts the wax onto the paper, which then pass through a fusing roller to fuse the wax onto the paper.

As the paper passes through the paper train, it first presses against the cyan color band of the ink transfer sheet. One or more heating elements arranged in a row on the thermal print head on the other side of the ink transfer sheet are turned on to melt small dots of the cyan dye. The melted dots are pressed against the paper. The paper continues moving through the paper train until it is partially ejected from the printer.

As the paper peels away from the ink transfer sheet, the unmelted cyan ink remains on the transfer sheet and the melted dye sticks to the paper.

Next, the color transfer sheet turns to expose the magenta band and the paper is pulled back into the printer, where it presses against the magenta band of the transfer sheet and the printing process is repeated. This process repeats itself for all of the colors used by the printer and then finally the page is completely ejected. At the end of this process the printed-paper will contain almost photographic quality color image.

1.3.2 Paper Handling

Any printer must have some mechanism to advance the paper line by line or page-by-page based on the printing method applied by the printer, once the printing of the line or page is finished. There are many different methods used for this purpose, some of them are useful to print single sheets some or more suited for big jobs where many thousands of sheets are to be printed continuously. Depending on the type of the job one wants to execute using the printer, one must carefully choose the printer and the paper handling procedure, because the output media that one can use with the printer depends on the this mechanism used for paper handling.

Next, we will see some of the common paper handling procedures offered by the printers.

1.3.3 Friction Feed

This is one of the oldest methods to move the paper through the printer. Typewriters use friction feed mechanism to feed the paper through its print mechanism. In this friction feed method the paper is squeezed through a large rubber roller called a platen and a smaller pressure roller. This type of paper feeding method is suitable only to print on the single sheet, which is fed manually into the printer.

The paper is additionally held around the platen above the area where the hammers strike it by a bail arm, which usually pivots out of the way when you load paper. While feeding the paper, with each sheet one has to pull out the bail arm, insert the sheet, properly align it, so that the print head doesn't type diagonally across the paper, lock the paper in-between the platen and the roller, bring the bail arm down to hold the paper close to the platen and finally signal to the machine that everything is ready for printing. This feeding method is not suitable to print a large job as one need to sit and manually feed the sheets one by one to the printer; on the positive side this paper feeding method is suitable for almost any kind of paper that can be fed into the printing mechanism. One can also print on different type of envelopes, cards etc. using this paper feeding method.

1.3.4 Automatic Feed

As the use of fraction feed printers increased, need for automatic sheet feeding increased and a device called automatic sheet feeder was introduced.

This device is also known as the bin-feed. As its name suggest one can keep a number of sheets inside this feeder and it will automatically feed them one by one to the printer. Unfortunately, these are one of the most complex and expensive device to add to a dot-matrix printer. Price rises with complexity and sheet feeders tend to be expensive. Most sheet feeders are designed for single-layer paper; i.e. you cannot put sheets with carbon paper into them and wait for the feeder to take care of that.

For copies, you must make two separate printings on separate sheets. Your printing time can run into double and triple when you need more than one copy. Laser printers have these cut-sheet feeders as a part of them. Because the Laser printer must be able to slip a single sheet from a stack, through the complex printing mechanism, and deliver it back to the user. Laser printers with significant differences in the paper handling facility are available.

One difference is in the capacity. Some Laser printers are made only for light, personal use has modestly sized paper bins that can hold about 50 sheets. That means every 10 to 15 minutes you need to load and remove the stack of paper. If you have a big print job, then this will be a big problem, as you have to sit full day changing the printer paper.

1.3.5 Rolling Feed

One way to reduce the number of times you have to feed a sheet of paper into the friction-feed mechanism is to make the paper longer. In fact, one can use one, long continuous sheet, some systems do exactly that, wrapping the long sheet around a roll. The printer just pulls the paper through, as it needs it. By firmly setting a roll-holder at the back of the printer, the paper can be kept in reasonable alignment so that the output will not get printed diagonally across the paper. The shortcoming of this system is that you will end up with one long sheet. You have to tear it to pieces or carefully cut it up when you want traditional 8.5 x 11-inch output.

1.3.6 Pin-feed and Tractor feed

Although roll-fed paper could be perforated at 11-inch intervals so that you could easily and neatly tear it apart, another problem arises. Most friction mechanisms are not perfect. The paper can slip so that, gradually, the page breaks in the image and the page breaks at the perforations no longer correspond. In effect, the paper and the image will not properly coincide. By locking perforations in the edge of the paper inside sprockets that prevent slipping, the image and paper breaks can be kept in perfect alignment.

Two different paper-feeding systems use sprocket paper to avoid slippage. Pin-feed uses drive sprockets, which are permanently affixed, to the edges of the platen roller.

Consequently, the pin-feed mechanism can handle only one width of paper, the width corresponding to the sprocket at the edges of the platen. Tractor-feed uses adjustable sprockets that can be moved closer together or farther apart to handle nearly any width paper that will fit through the printer. Sprocket tractor feed is available as

unidirectional tractor and bi-directional tractor. As the names imply, a unidirectional tractor only pulls (or pushes) the paper through, in a single direction normally forward.

The bi-directional tractor allows both forward and backward paper motion, which is often helpful for graphics, special text functions (printing exponents, for instance), and lining up the top of the paper with the top of the print head i.e. for proper automatic paper feeding.

1.3.7 Push and Pull Tractors

The original tractor mechanism for printers had two set of sprockets, one to feed the paper into the printer and another to pull the paper out of the printer. Later it was found that one set of tractor is more than enough to properly lock the printed image in synchronization with the paper. A single set of sprockets can be located in one of two positions, either before or after the paper wraps around the platen in front of the print head. Some of the printers allow a single set of tractor to be used in any of these positions, whereas most of the printers have a fixed tractor location.

The push tractor is placed in front of the printer in the paper path, before the paper enters the printer. As the name suggest it pushes the paper into the printer. The pull tractor is located into the paper path after the paper comes out of the printing mechanism; here again as the name suggest the pull tractor pull the paper across the platen through the print mechanism. This pull mechanism is simpler to implement and there is less of a chance of paper jam using this method compared to the push tractor method.

One advantage with the push tractor is one can remove the last sheet from the printer without wasting an extra sheet through the printer. In the case of pull tractor one need to eject an additional page to remove the last printed sheet from the printer. The push type of tractor can also be used to pull the paper backward with relative ease compared to the pull type of tractor.

Most of the pull tractor printers can operate only in one direction i.e. one cannot use them to push the paper backward. As the pull type tractors are more suited for the high speed operation, waste of a single sheet or not being able to push the paper backward does not make much difference on the use of pull type of tractors in high speed printers.

1.3.8 Paper Control

Printers differ to a great degree in how precisely they can move paper through their mechanisms. Some are designed to allow the sheet to be moved in increments of the tiniest fractions of an inch (as small as 1/216th inch). Most of the printer allows you to change the line spacing from 6 lines per inch for manuscripts to 8 per inch for business letters, you can also add more accuracy in printing graphics by varying the inter line spacing. Printers also vary in the control they provide in the horizontal direction, motion across the paper.

Most of the printer allow the character pitch i.e. width of the character (number of characters per inch) in the text mode and the spacing of dots in graphics modes to be changed. This versatility is necessary for the printing of proportionally spaced text and for printing in multiple graphic densities.

1.4 Summary

In this unit, we have discussed about printer, which is only responsible to produce output in the form of hardcopy. We have also described different type of printer used in computer field. We have introduced working process of different printers. Finally, we have explained about paper handling of different printers.

1.5 Model Questions:

Write a short note on the following.

1. What are roles of printer? Can you explain impact type of printers used in computer?

2. How non-impact printer is suitable? Explain its different advantage.

3. Describe the working process of different printers.

UNIT-2: INTERFACING AND INSTALLING THE PRINTERS

Structure:

2.0 Objectives

2.1 Introduction

2.2 Printer Interfacing Software Tools

2.2.1 Plain ASCII Standard Data

2.2.2 Control Code/Escape Code

2.2.3 Bit-Image Data

2.2.4 Page Description Language (PDL)

2.2.5 PostScript

2.3 Printer Interfacing Hardware Tools

2.3.1 Serial Port

2.3.2 Serial Port Connectors

2.3.3 Parallel Port

2.3.4 Uni & Bi-Directional Parallel Port

2.3.4.1 Enhanced Parallel Port (EPP)

2.3.4.2 Extended Capabilities Port (ECP)

2.3.5 Parallel Port Connectors

2.3.6 USB (Universal Serial Bus)

2.3.6.1 USB Cable/Connector

2.3.6.2 USB Troubleshooting

2.3.7 Infrared Connecting Device

2.3.7.1 Installation

2.3.7.2 Infrared Device Troubleshooting

2.4 Installing and Configuring Printers in Windows

2.5 Summary

2.6 Model Question:

2.0 OBJECTIVES

At end of this unit, student will be able to:

- Understand about the basics of interfacing and installing different type of printers.
- Explain the installation of printers with the help of software.
- Distinguish between serial port and parallel port of printers.
- Know the concept various type of interfacing devices used to connect printer.
- Analyze statistical feature for suitability of USB connecting device.
- Generate awareness about working process Infrared connector.
- Identify the various clue to shoot the troubles occurs in different type of ports.

2.1 INTRODUCTION

In the computer, it is necessary to interface printer with CPU properly. So in this regard, we have check configuration of interfacing device before attaching with the CPU. When configuration matched, then try to installs the device and takes the operation.

In this unit, we look at how the information is transferred from computer to the printer.

2.2 PRINTER INTERFACING SOFTWARE TOOLS

The printer is connected to the computer through a data cable, the printer can interpret any information send by the computer to the printer in different ways, based on this interpretation, and printer acts differently on the received data.

Information send by the computer to the printer can be interpreted as:

- Plain ASCII Data
- Printer Control Code/Escape Code
- Graphic/Bit-Image Data
- Page Description Language Statement

2.2.1 Plain ASCII Standard Data

The plain ASCII data is a byte of information sends from the computer to the printer, it contains ASCII (American Standard Code for Information Interchange) code of the character to be printed; it is up to the printer to convert this code into proper character shape and print it.

When a printer receives an ASCII code, it consults its own memory for the shape of the received code, for example, when the code 65 is received, the printer searches in its own memory for the shape of the character 65. In standard ASCII, code 65 is used for the character "A", so the printer will get the bitmap image of the character "A" from its

memory and send it to the printing mechanism. The printing mechanism, based on this information, will print the shape "A" on the paper.

2.2.2 Control Code/Escape Code

ASCII codes are also used to inform the printer about any special function such as printing bold letters, printing underline etc. required by the user. ASCII codes that are used to communicate these special functions to the printer are called printer control codes. These codes are also referred as embedded commands, as these codes are embedded in the data coming from the computer.

Some of the common function such as backspace, line feed, form feed etc. are included in the standard ASCII character set itself. These functions are assigned specific character value. To linefeed the printer, computer can just send the ASCII code 10 to the printer. The character 10 is the linefeed character in the standard ASCII code. Receiving this character, the printer will move the paper one line up. The entire group of these special ASCII codes is termed as control characters. The number of ASCII characters available for the printer commands are very few and the number of functions that the printer can carry out are many. To allow these additional instructions to be sent to the printer most printers use special combination of characters as printer control codes. When a combination of ASCII code is used as control code in most of these codes, the first character is always ASCII code 27 (01BH), or the Escape character. As most of these control codes start with the escape character, these codes are also known as Escape control codes. In most of these commands, the escape character by itself does nothing. It is only used to inform the printer that the characters that follow the Esc code should be interpreted as control codes and should not be printed as plain ASCII data. Most of the printers use Epson or IBM compatible control codes, so these codes have become an industry standard in the Dot-Matrix printer segment. Laser and Inkjet printers use the Hewlett Packard codes as standard.

2.2.3 Bit-Image Data

Sometimes instead of sending ASCII code or control code to the printer, the computer may want to directly control the printing of the image or control the print head of the printer. In this situation computer sends the complete bit-image of the image to the printer. If the printer is a dot-matrix printer, it will generate the image by firing its pins according to the bit-image data being received. Before sending this type of data to the printer, most of the printer requires some type of control code to be sent from the computer, to inform the printer that the data being sent by the printer is not ASCII code, instead it is bit-image data, and it should be properly interpreted and printed by the printer.

2.2.4 Page Description Language (PDL)

Some printers are especially the page printers such as Laser printer can generate image according to the statement of page description language sent by the computer. These languages can be used to control any of the printer supporting them.

The major advantage of languages is that they are hardware independent, i.e. when a computer sends output to the printer in PDL format, the printout can be taken on any printer that supports these languages. Most of the Page Description Language resembles a high-level language such as Pascal or C, with flow control and logic statements etc.

When one uses page description language to be sent output to the printer it will be the job of the printer to generate the image to be printed based on the instructions received. This makes it necessary for the printers supporting PDL to have a good deal of computing power and memory capacity, otherwise the printing process by these printers will be very slow.

There are two very common industry standards PDL are PostScript and PCL.

2.2.5 PostScript

PostScript is a page description language used to define text, graphics and complete page layout in a high-level language like format. This language was developed by the Adobe systems.

One major advantage of the PostScript is that, it is device independent i.e. the same PostScript code when send to different devices produce the same type of image using the maximum capability such as the resolution etc. available on that particular device. One does not require a different PostScript code for different device. For example, when a PostScript file is send to 300-DPI printer, a 300-DPI image is printed, the same file when send to a 1200-DPI PostScript compatible typesetter will print the at 1200-DPI resolution.

A printer that has PostScript facility require lot of computing power and memory because the PostScript language is very complex and computer intensive and it is the job of the printer to convert the incoming PostScript instructions into proper image.

2.2.6 Printer Control Language (PCL)

PCL or the Printer Control Language was introduced by the Hewlett Packard to control their inkjet printers. This control language was later expanded as PCL3 to control their first laser printer, the original LaserJet printer. PCL is not a true page description language like the PostScript; this is more of an expanded printer control command set. To control the printer using the PCL, one needs to send long strings of control codes. Even the PCL5 introduced in 1990 is not a device independent language like the PostScript; It currently works only with laser printers, so its code cannot be used to drive some other device such as a typesetter. Although, the PCL5 lack many features of the PostScript language, it can produce PostScript quality image on the paper. HP has currently introduced the PCL6, which has improved the printing speed and provided many new routines to print the scanned images, faster and better.

2.3 PRINTER INTERFACING HARDWARE TOOL

After looking at how the data sent from the computer is interpreted by the printer, let us now see how the printer can be physically connected to a computer. Printer can be connected to the computer at Serial, Parallel, USB or IrDA (Infrared) port.

2.3.1 Serial Port

The computer to communicate with a device, one bit at a time, uses the serial port. Data moves serially one bit at a time from one device to another. To communicate using the serial communication, two communicating device must use the same communication parameters such as the communication mode, baud rate (bits per second), parity etc.

If the serial cable for your device contains 25-pin connector and the serial port on the computer has only 9-pin connector, then many 25-pins to 9-pin converters are available in the market and we are capable for connecting 9-pin device to a 25-pin port connector using a 9-pin to 25-pin converter.

2.3.3 Parallel Port

A parallel port is a fast and trouble free connection you can have on your computer. This port is also known as the printer port. It is commonly used for the printer interfacing. Some peripherals such as the scanner can also use this port to communicate with the computer. In this interfacing system, the signals travel through data cable, eight bits at a time. Other than the data, some control signals are also passed through the cable between the printer and the computer. The parallel interface is one of the simplest communication systems available. In these eight different lines are provided for the eight data lines, for each signal required to control the communication process a new wire is provided. This makes the circuitry required to drive the parallel signals very simple. However this simplicity has one drawback that is a cable is required for the transmission. The increased numbers of lines make the cable required for the parallel transmission very thick. The numbers of lines going together have one more problem of crosstalk. When the electrical signals travel together in a cable, these signals start to affect each other as they travel down the cable. This makes the signal from one wire to change the signal on the other wire or the problem of cross talk (signal crosses over and start talking with other signals) appears.

To prevent this problem of cross talk it is recommended to keep the length of the parallel cable limited to maximum 10 feet. If the cable length is more than this, there is a chance of cross talk appearing in the line.

2.3.4 Uni & Bi-Directional Parallel Port

The original parallel interface provided on the initial IBM computer's were designed as printer output port so they were uni-directional i.e. they could only send the information to the device connected to the port but they could not receive any information from the device.

Later IBM made some changes in the design and made the parallel ports capable of bi-directional operation.

2.3.4.1 Enhanced Parallel Port (EPP)

Enhanced Parallel Port or EPP is a new parallel port specification that allows high-speed two-way communication between the computer and the device, other than a printer or scanner.

2.3.4.2 Extended Capabilities Port (ECP)

Extended Capabilities Port or ECP is also a new high-speed two-way communication parallel port specification. ECP port is mainly used for a printer or scanner. To connect to a printer to the ECP port you would require

- EPP/ ECP port
- ECP capable printer

- EPP/ECP capable printer cable
- Windows 98/2000/ME/XP operating system

Parallel port provided on the motherboard can be setup as EPP/ECP port using the following procedure

- Enter BIOS setup during computer power on by pressing Del key.
- In the BIOS setup main menu, select Peripheral Setup option.
- Look for Parallel Port Mode.
- Change the mode to EPP/ECP.
- Save and Exit BIOS setup and restart the system.

2.3.5 Parallel Port Connectors

IBM has followed the standard developed by the printer manufacturer Centronics when deciding the design of the parallel interface. Although IBM followed the signals of the Centronics, it did not approve of the 36-conductor Amphenol connector used by the Centronics, instead IBM selected 25 pin D-shell subminiature connector for their parallel connection. Since then, the printer cable for the parallel printer have a 25-pin D-shell connector on one side which is connected to the computer and the other side of the cable will have 36-conductor Amphenol connector, to connect to the printer.

2.3.6 USB (Universal Serial Bus)

The Universal Serial Bus (USB) makes your computer easier to use with advanced plug-and-play capabilities. USB is not a connection only for printers; USB is a bus standard that can be used to connect various external hardware devices, such as keyboards, mouse, speakers, and digital cameras to the computer.

USB devices are hot pluggable, i.e. you can add devices to your computer easily without having to shutdown the computer. External devices can be connected at any time, even when the computer is running. Because of the Plug and Play of USB capability, computer can identify USB devices when they are plugged in and automatically configure them by loading the appropriate driver. USB devices can be powered by drawing current directly from the USB cable. If you have a new computer, you will find four/six USB ports on the motherboard. If you have an old system, you can connect a PCI USB adapter card, to get USB facility.

In addition, one needs to have the latest operating system for the USB support, currently Windows 98/ME/XP/2000 and later versions of Windows operating system support USB port. Older Windows-3.1/95 etc. does not support USB. USB port has option to daisy chain the devices connected at its output socket. One can connect a maximum of 127 peripheral devices on a single USB port. Each USB port connects to a single device; to connect multiple devices one can use hubs. One can connect one hub into other to provide several jacks to connect more devices to a single USB port. By connecting hubs together, up to 127 total devices, including hubs can be attached to the host computer. Each device can be located up to five meters from its hub.

2.3.7.1 USB Cable/Connector

The USB specification defines a standard connector, socket, and cable, which all USB devices can use. The USB hub uses a type A connector, and the device uses a type B connector.

- **Type A:** Connector/jack is the thin wide slots provided at the back of the computer to connect a USB device.
- **Type B:** Connector/jack is a small, nearly square hole provided on them USB devices (printer, scanner etc.), into which the mating B plug is connected.

All USB cables will have A plug at one end and B plug at the other end, one end attached to the hub and the other end attached to the device. There are two types of USB cables: high speed and low speed. Low-speed cables differ primarily in their shielding.

Plugging a high-speed device into a low-speed cable can cause signal distortion over long distances. The maximum allowed length of a USB cable is five meters. Both high speed and low speed USB cable use a four-wire cable. Two wires (+Data and -Data) transfer data as a differential digital data signal, i.e. The voltage on the two wires is of equal magnitude but opposite polarity. When this data signal is subtracted from the other, results in any noise in the signal getting canceled. One wire is used as power line, which contains a five-volt DC voltage. This power line is used to provide power for devices connected to the USB cable. Final wire is used as ground return line.

On the high-speed line, the two data lines are twisted together. Following table has shown the color-coding scheme that is used in the USB cable.

Color	Signal
Green	+Data
White	-Data
Red	+5Vcc
Black	Ground

2.3.7.2 USB Troubleshooting

Check the following to troubleshoot printer connected to the USB port.

2.3.7.3 Check Window Version

When using USB port to connect a printer, make sure that the operating system supports the USB port. In Windows 3.1, Windows 95 or Windows NT, the USB port will not works. Operating systems Windows 98 onward provide support for USB connection. If your system has Windows XP or Windows 2000 then the USB port should work without any problem. If your computer system uses some older version of operating system then upgrade to a new version before installing the USB printer.

2.3.7.4 Check USB Controller

Some of the initial motherboards with USB support do not provide proper support for the USB port. Therefore, if you have one of those motherboards you need to upgrade the motherboard with a new one.

To check the version of USB controller

- Click on Start, Settings and Control Panel.
- Double Click on System.
- Click on Device Manager Tab.
- Double Click on Universal Serial Bus Controllers.
- Double Click on the Intel 82371 USB PCI to USB Universal Host Controller entry. If the controller name being shown is different ways then you have a new controller, you can skip the next step.
- In the General tab check the hardware version, it should not be 000.

If it is 000 then you need to upgrade your motherboard or disable USB on the motherboard and buy a new PCI USB card.

2.3.7.5 Read the manual before installing USB printer

Installing a USB printer is not as simple as installing a serial or parallel printer, some printer require that the printer software should be installed before connecting the printer, some require that you connect the printer halfway between installing the software.

If you do not follow these steps exactly, then the printer may not work properly, or may not work at all.

If you make a mistake at the first install then before reinstalling the printer, you cannot just reinstall to correct the problem; you need first remove the files from the first install.

This again is not as easy as just uninstalling the printer from control panel, you can find detailed uninstall procedure in the printer manual or in the printer manufacturers web site.

2.3.7.6 Printer features are missing

After installing a printer, if you find that many of the printer's features are missing in the property dialog box of printer then you may be using some old version of the printer driver. Reinstall the printer with latest driver.

2.3.7 Infrared Connecting Device

IrDA (Infrared Data Association) is an international organization that creates and promotes interoperable, low cost, infrared data interconnection standards.

However, the IrDA mean a group of short range, high speed and bidirectional wireless infrared port, defined by the IrDA group. IrDA port allows a variety of devices to communicate with each other. Cameras, printers, portable computers, desktop computers, personal digital assistants (PDAs), and mobile phones can communicate with compatible devices using this technology. IrDA port supports transmission of data

between two devices over short-range at speeds between 9.6 kilobits per second (Kbps) and 4Mbps. IrDA is a small semitransparent red window you may see on your notebook computer, PDA, printers or even on the mobile phones.

2.3.7.1 Installation of Infrared Device

This section describes how to install infrared devices. If you change the infrared adapter model connected to the computer, you should reinstall it. The procedure for installing an infrared device depends on whether it is a Plug and Play device or a non-Plug and Play device.

To install a Plug and Play IrDA device

If it is an external device, attach it. The Infrared software loads automatically.

Double-click the Infrared icon in Control Panel to activate the infrared device. If there is no Infrared icon in Control Panel, press F5 to make the icon appear.

To install a non-Plug and Play IrDA device

If it is an external device, attach it. In Control Panel, double-click the Add New Hardware wizard. Follow the instructions on the screen, making the following choices:

If asked whether the device you want is in the list, click no. When asked whether you want Windows to search for your new hardware, click no. When the Hardware Types box appears, click Infrared devices. When asked to select a device, accept the default selection Generic Infrared Serial Port or dongle. (The other devices listed are Plug and Play devices.) When asked for such information as the communications port that the infrared device is physically connected to, if you are not sure, accept the defaults.

Using the Infrared Monitor

You can use the Infrared Monitor to keep track of your computer's infrared activity. You can activate Infrared Monitor once your infrared device is installed by double-clicking the Infrared icon in Control Panel or by clicking the animated Infrared icon in the system tray on the Taskbar.

Printing with Infrared Port

Before attempting to print, make sure the infrared communications driver is properly installed and the infrared devices are enabled. Make sure the infrared communications driver is installed on the computer. Bring an infrared-enabled printer within range. Your computer might automatically detect and install the printer. If not, make sure the printer driver for the infrared-capable printer is installed on the computer.

Try the Print option in a program. You can make printers without built-in infrared ports infrared-capable by connecting an infrared adapter made for printers into the printer's parallel port. If a parallel cable is also used to connect the computer to the infrared printer adapter, you can use either the infrared link or the parallel cable to print. The infrared link is used when you select the virtual parallel port; the cable is used when you select the physical parallel port. To check the infrared link to the printer, make sure the correct printer driver is installed for the infrared-capable printer. (Most printers with

built-in infrared ports are Plug and Play devices, which are installed automatically). Then use a program to print over the infrared link.

If the program prints on an infrared-capable printer, it confirms that the infrared driver is properly installed.

2.3.7.2 Infrared Device Troubleshooting

Make sure that the IrDA port on the computer and the printer are clean and unblocked. Move the computer/printer to one side to bring them in proper line of sight. Check that nothing is blocking the infrared ray moving between computer and printer. Make sure that the infrared port on the computer is active. Check the BIOS setup. Check that no other infrared emitting device is interfering with the communication, such as a TV remote handset.

Printer Sharing

In an organization, one does not need a printer for each computer. To save the cost of the additional printer one can share the printer between two or more computers. Several printer-sharing devices are available in the market, some of them are hardware based and some are software based.

Sharer Switch

This is also called Printer Sharer switch or an A/B Switch. This device is most commonly used to share the printer between two computers; one can also connect more than two computers to one printer using this type of switch. This device works like a simple multiple switch, when the switch is in A position the information from the computer A is routed to the printer and when the switch is changed to position B then the output from the computer B is send to the printer. This process is almost equal to removing the printer data cable from the computer A and connecting it to the computer B, only difference is one is saved the work of removing and reconnecting the cables every time another computer need the printer. This type of printer sharer is available for both serial as well as parallel interfaces. One can use them to share one printer with 2, 3 or 4 computers. Some of these sharers can do auto switching i.e. they can sense the computer from which the data is coming and can automatically switch to it.

Sharing through Network

If your computer is connected to a network, most of the networks provide printer-sharing facility.

If your computer is connected to the local area network and you are using Windows 98 then follow the below given steps to share your printer

- Click Start, point to Settings, and then click Printers.
- Click the printer you want to share.
- On the File menu, click Properties.

- Click the Sharing tab, and then Click Shared As, type a name you want the others to see for your printer.

You can only share a printer that is connected to your computer

If the Sharing tab is not visible, you need to enable file and prints sharing services as explained below.

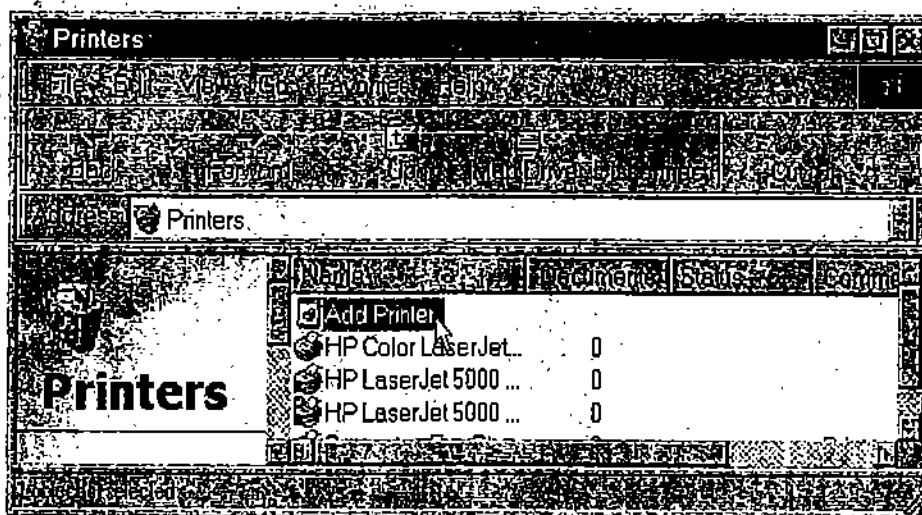
- Open the Network dialog box by clicking Start, pointing to Settings, and clicking on Control Panel
- Next, double-click on Network.
- Click "File and Print Sharing".
- Select the check box for the sharing options you want.
- A check mark indicates the feature is activated.
- Click OK.

2.4 Installing and Configuring Printers in Windows

You can install printers in Windows using the Add Printer Wizard.

Using the Add Printer Wizard

Windows provides the "Add Printer Wizard" to simplify installing printers, and a central place - the Printers folder - for running the wizard and for managing printing processes.



You can open the Printers folder in the following ways

- From the Start menu, point to Settings, and then click on Printers or
- In My Computer, double-click Printers or
- In Control Panel, double-click Printers.
- In the Printers folder, double-click Add Printer.

- The Add Printer Wizard leads you through the process of setting up and configuring a printer.

The only difference between installing a network printer and a local printer with the Add Printer Wizard is that you must specify the path to the network printer.

2.5 Summary

In this unit, we have discussed about installation and configuration of different printers. We have also described different type of printer connecting cables used in computer field. We have introduced the installation of printer using Windows' Add Printer Wizard. Finally, we have explained about installation of USB and Infrared cable and their trouble shootings.

2.6 Model Question:

Write a short note on the following.

1. What are roles of software tool during installation of printer? Can you explain different type of software approach to synchronized printers with computers?

2. What do you understand by parallel port connection? Explain how it is better as compared to serial port connection.

3. Describe the trouble shooting of USB and Infrared cables.

UNIT-3: SERVICING AND TROUBLE-SHOOTING THE PRINTERS

Structure:

- 3.0 Objectives
- 3.1 Introduction
- 3.2 Prevention in Servicing
- 3.3 Troubleshooting Dot-Matrix Printer
 - 3.3.1 Print Quality Poor
 - 3.3.2 No Power to Printer
 - 3.3.3 Power OK but No Printing
 - 3.3.4 Grinding Noises
 - 3.3.5 Paper Jamming/Misfeeding
- 3.4 Troubleshooting of Inkjet Printers
 - 3.4.1 Print Quality Poor
 - 3.4.2 No Power to Printer
 - 3.4.3 Power OK but No Printing
 - 3.4.4 Grinding Noises
 - 3.4.5 Paper Jamming/Misfeeding
- 3.5 Troubleshooting Laser Printers
 - 3.5.1 Print Quality Poor
 - 3.5.2 No Power to Printer
 - 3.5.3 Grinding Noises
 - 3.5.5 Fuser Malfunction
- 3.6 Summary
- 3.7 Model Questions

3.0 OBJECTIVES

At end of this unit, you will be able to:

- Understand about the basics of trouble shooting of the printer.
- Explain the trouble shooting of the Dot matrix printer.
- Know the concept of various types of trouble shooting of the Ink jet printer.
- Analyze statistical prevention in servicing of a printer.
- Generate awareness about trouble shooting of the Laser printer.

3.1 Introduction

Ninety percent of the printer problems can be easily rectified by a non-technical person, because most of the time, the problem does not exist inside the printer. Instead, most of the faults are reported about the printer side. Due to some mechanical fault or some fault in the printer cable, it could also be due to printer not being setup properly. It could be due to some problem in the software.

In this unit, we shall look at some common problem faced when using a Dot-Matrix printer, Inkjet printer and Laser printer.

3.2 Preventions in Servicing

Before starting on the troubleshooting let, we first see some of the preventive maintenance steps that one should take to keep the printer in good working condition.

When one opens a Dot-Matrix printer, one may find many small pieces of paper inside the printer. One should periodically vacuum out these paper scraps. Clean the printer paper path, ribbon path and head assembly movement path with a dry soft cloth. You can also use a soft brush for this purpose. Also clean the outer surface of the printer with some soft cloth, damp with warm soap water, do not use cloth that is dripping, and instead use a damp cloth.

If you print on self, sticking labels and envelopes with gum, then the printer rollers may have attracted some glue from these materials. Remove this glue buildup by using alcohol and a clean cloth. If you periodically clean and lubricate your printer, you can avoid expensive repairs later. Once the cleaning is done, check for any loose belts etc. If any belt is loose then replace it. Next comes lubricating. Lubricating of the head assembly movement mechanism can be done periodically to have a smooth movement of the print head.

When applying lubricants one should be careful that it should not be applied to wrong places, otherwise printer damage can occur. Whenever some lubricant is used, it should be Silicone or Teflon based. One should apply just enough lubricant so that the extra lubricant will not accumulate dust and create problem for the printer. One can safely lubricate most of the mechanical parts, for smooth working of the printer.

For long life of the printer's print head, keep the print head as cool as possible, do not **stack up** thing on and around the printer that may stop proper airflow. Also, for the print head's life you should not use cheap ink to re-ink the ribbon cartridge. The cheap ink used for the re-inking contains acid which corrodes the print head pins and as some necessary lubricant provided in more expensive inks is missing **in these inks**, after some time, the print head pins start to stick. One should not turn the printer platen roller or ribbon carriage assembly motor, which is moved by another stepper motor.

If your printer has, friction feed and tractor feed option then do not engage **both the** friction feed and tractor feeds at the same time. If you engage both the tractor feed and the friction feed at the same time then the paper feeding will not be proper and the paper may get stuck in the printer.

3.3 TROUBLESHOOTING DOT-MATRIX PRINTER

3.3.1 Print Quality Poor

Ribbon worn out or Damaged: Remove the ribbon and check to see if the ribbon appears to be dried out, damaged or rotation has seized. Replace the ribbon with the genuine item.

Print Head Too Close to the Platen: Adjust the print head lever to ensure that the print head is not too close to the platen and hence paper. This will cause not only smudging but may lead to paper jamming and in some cases, print head damage.

Ribbon Drive Assembly Faulty: Print a self-test, or send information to be printed. While this is happening, observe that the ribbon is turning in both directions. If the ribbon is not turning or only turning in one direction then the ribbon drive assembly could be at fault and require replacement

Print head is Faulty: Remove the print head. Visually inspect the pins. If some are missing or damaged then the print head is faulty. Depending upon the brand of your printer, print heads can be repaired for a fraction of the cost of a replacement print head. Working on print heads is as intricate as working on a watch, be careful when trying to repair.

3.3.2 No Power to Printer

Fuse has Blown: Check the fuse with a meter and replace the fuse. Ensure the correct voltage and amperage.

Faulty Power On-Off Switch: Check with a meter and replace the switch. Ensure the correct voltage and amperage.

Internal Power Supply Failure: It is most common fault. Repair yourself only if you know what you are doing. There may be chance of multiple damage can be

cure through replacing one or two components will only allow power to circulate temporarily causing greater damage.

Main Logic Board is Shutting down the Power Supply: Repair the logic on the main logic board and in some cases when damage is too extensive or parts are unavailable (custom chipsets), the main logic board will require replacement.

3.3.3 Power OK but No Printing

Printer Cable Not Installed Properly: Ensure that both ends of the cable are correctly installed.

Printer Cable is Faulty: Check for possible pin damage and replace if necessary.

No Paper or Paper Not Installed Properly: No paper will result with the printer going into alarm, not coming on-line or ready. Install some paper. If there is paper, ensure that, it is correctly installed. Tractor fed paper needs to be properly aligned and driven up to the start of the platen roller. Ensure that the on-line button has been selected. There are sensors along the paper path; if these are defective, they will cause the printer not to print. Check the sensors along the paper path and replace if necessary.

Printer Port Connection at PC is Faulty: Try printing plain text via DOS (you can even use the print screen command on your keyboard while you are in the DOS environment). If still not printing, refer to your printer's owner manual and perform a self-test of your printer. If the printer prints OK, the fault is more likely with the computer or the printer cable, not with the printer.

Self-Test is OK and PC Printer Port is OK: Faulty main logic board. Repair input/output circuit on main logic board. Extensive damage may require replacing the main logic board.

3.3.4 Grinding Noises

Obstruction along Paper Path: There might be something caught along the paper path or blocking the print head from correctly moving from side to side.

Ribbon Drive Assembly Faulty: If the noise appears to be coming from the ribbon drive assembly, stop the printing, remove the ribbon drive assembly and visually inspect for worn or broken gearing. Replace if necessary.

Ribbon Carriage Assembly Dirty: Inspect the carriage assembly, remove any dust or paper particles. Lubricate where necessary.

Gear Train Assembly Worn Out or Damaged: Inspect the gear train assembly. If worn out or damaged, replace the gear train assembly.

3.3.5 Paper Jamming/Misfeeding

Paper Quality Poor: Ensure that the paper you introduce into your printer meets with the standards of the printer manufacturer and replace if necessary.

Obstruction along Paper Path: There might be something caught along the paper path.

Sensors Faulty: Faulty sensors along the paper path will not provide the printer with the correct information about the paper's position in the printer and hence the printer will fall into a false paper jam mode. Inspect and test the sensors and replace if necessary.

Tractors Worn Out: For printers that use tractor fed paper. The tractors may worn out and can be out of alignment with one another. Observe the tractors as they feed paper. A good way to test this is to form feed multiple sheets of paper and see what happens. If the tractors require replacement, always replace both the LHS and RHS tractors at the same time.

Main Logic Board is Faulty: In some cases, the main logic board could be the cause of a false paper jam or alarm mode. Repair or replace the main logic board.

3.4 TROUBLESHOOTING OF INKJET PRINTER

3.4.1 Print Quality Poor

Ink Cartridge Empty or Faulty: Remove the ink cartridge and check to see if the ink cartridge appears to be dried out or damaged. Replace the cartridge with the genuine item. It is sometimes difficult to tell whether the ink cartridge is empty, as the inkpad still may appear to be damp.

Print Head Faulty: Remove the print head and inspect for blockages. Replace if necessary.

Purge Unit Assembly Faulty: If your purge unit assembly is damaged, it will cause your ink cartridges to dry up and render them faulty. Inspect the operation of the purge unit assembly to ensure that the print head is parking itself properly thus sealing itself from the air while not in use. Replace the purge unit assembly if necessary.

Printer Requires Service: A build up ink, paper dust and residue may occur, which will require the printer to be totally serviced, stripped and cleaned. An alignment of the print head, purge unit and gearing will be required.

Printer Operation Out of Synchronization: You may have experienced a paper jam and in attempting to retrieve the paper have thrown gearing out of synchronization. This will result in the gearing, print head and purge unit assembly to become misaligned. Dismantle the printer, realign the gearing,

print head and purge unit assembly. Service and clean the printer while you have it dismantled.

3.4.2 No Power to Printer

Fuse Has Blown: Check the fuse with a meter and replace the fuse. Ensure the correct voltage and amperage.

Faulty Power On-Off Switch: Check the switch with a meter and replace the switch. Ensure the correct voltage and amperage.

Internal Power Supply Failure: Repair yourself only if you know what you are doing.

AC Power Supply has Failed: Most AC adapters are sealed, repair yourself only if you know what you are doing. There are chances may multiple damage, replacing one or two components will only allow power to circulate temporarily causing greater damage.

Main Logic Board is Shutting Down the Power Supply: Repair the logic on the main logic board and in some cases, when damage is too extensive or parts are unavailable (custom chipsets), the main logic board will require replacement.

3.4.3 Power OK But No Printing

Printer Cable is not installed properly: Ensure that both ends of the cable are correctly installed.

Printer Cable is Faulty: Check for possible pin damage and replace if necessary.

No Paper or Paper Not Installed Properly: No paper will result with the printer going into alarm, not coming on-line or ready. Install some paper. If there is paper, ensure it is correctly installed. Ensure that the on-line button has been selected. There are sensors along the paper path; if these are defective, they will cause the printer not to print. Check the sensors along the paper path and replace if necessary.

Printer Port Connection at PC is Faulty: Try printing plain text via DOS (you can even use the print screen command on your keyboard while you are in the DOS environment). If still not printing, refer to your printer's own manual and perform a self-test of your printer. If the printer prints OK, the fault is more that likely going to be with the printer cable or computer, not with the printer.

Self-Test is OK and PC Printer Port is OK: Faulty main logic board. Repair input/output circuit on main logic board. Extensive damage may require that the main logic board to be replaced.

3.4.4 Grinding Noises

Obstruction along Paper Path: There might be something caught along the paper path or blocking the print head from correctly moving from side to side.

Purge Unit Assembly Faulty: If the noise appears to be coming from the purge unit assembly, halt the printing, remove the purge unit assembly and visually inspect for worn or broken parts. Replace if necessary.

Gear Train Assembly Worn Out or Damaged: Inspect the gear train assembly. If worn out or damaged, replace the gear train assembly.

Printer Operation Out of Sync: You may have experienced a paper jam and in attempting to retrieve the paper have thrown gearing out of sync. This will result in the gearing, print head and purge unit assembly to become misaligned. Dismantle the printer; realign the gearing, print head and clear unit assembly. Service and clean the printer while you have it dismantled.

3.4.5 Paper Jamming/Misfeeding

Paper Quality Poor: Ensure that the paper you introduce into your printer meets with the standards of the printer manufacturer and replace if necessary.

Obstruction along Paper Path: There might be something caught along the paper path.

Sensors Faulty: Faulty sensors along the paper path will not provide the printer with the correct information about the paper's position in the printer and hence the printer will fall into a false paper jam mode. Inspect and test the sensors and replace if necessary.

Paper Pickup Delivery Assembly Faulty: The assembly of paper pickup delivery may be dirty or worn out. Remove the assembly of paper pickup delivery and clean the rollers. Reinstall and test. If the problem reoccurs, replace the assembly.

Main Logic Board is Faulty: In some cases, the main logic board could be the cause of going into a false paper jam and alarm mode. Repair or replace the main logic board.

3.5 TROUBLESHOOTING LASER PRINTERS

3.5.1 Print Quality Poor

Toner Cartridge Faulty: Remove the toner cartridge and check to see if the toner cartridge appears to be leaking, damaged or rotation has seized. Replace the toner cartridge.

Drum Unit Faulty: Some LaserJet printers have a separate drum unit, remove the drum unit being careful not to allow direct sunlight or continuous exposure to

daylight and do not touch it even if it appears to have residue of toner on it. It is probably faulty anyway but if it is not, then touching it will damage it.

Transfer Corona Assembly Faulty: Depending on the brand of your LaserJet, this item will be either a roller assembly or a wire assembly. Remove, clean thoroughly, replace and test. If yours is the roller assembly type, make sure that you do not touch the roller; the oil on your skin will damage it.

Fuser Unit Assembly Faulty: Probably, the most common places may cause poor print quality and paper jamming. Inspect the fuser unit assembly for scratches on the fuser roller. Ensure that the printer has turned off for at least one hour, as extreme temperatures exist within this item. If scratches are in evidence, replace the entire assembly, you can be sure that the thermostats, thermo switches, gearing and lamp will fail next. It is always better to replace the entire assembly.

High Voltage Power Supply Faulty: Classic signs are blank pages, black pages or faint print. Replace the high voltage power supply.

Laser Scanner Unit Assembly & Optical Fiber Cable: Missing print, faint print. Try cleaning the mirror, the front of the scanner unit assembly and then test it. Most of the time, this procedure should work. If this procedure fails to rectify the fault, replace the scanner unit assembly. It is also good practice to replace the optical fiber cable as well.

3.5.2 No Power to Printer

Fuse Has Blown: Check the fuse with a meter and replace the fuse. Ensure the correct voltage and amperage.

Faulty Power On-Off Switch: Check the switch with a meter and replace the switch. Ensure the correct voltage and amperage.

AC Power Supply Failure: Replace the AC power supply. Try to repair only if you are sure about what you are doing. The chance may be there of multiple-damage, replacing one or two components will only allow power to circulate temporarily causing greater damage.

PCA Interface/Main Logic Board is Shutting down the Power Supply: Repair the logic on the PCA interface/main logic board and in some cases, when damage is too extensive or parts are unavailable (custom chipsets), the PCA interface/main logic board will require replacement.

DC Controller is Shutting Down the Power Supply: Repair the logic on the DC controller and in some cases, when damage is too extensive or parts are unavailable (custom chipsets), the DC controller will require replacement.

DC Power Supply is Faulty: Repair or replace the DC power supply. DC power supply feeds the DC controller. Quite often if this is the culprit, then damage to the DC controller could very well have occurred.

Power OK but no Printing: Printer Cable not installed properly. Ensure that both ends of the cable are correctly installed.

Printer Cable is Faulty: Check for possible pin damage and replace if necessary,

No Paper or Paper not installed properly: No paper will result with the printer going into alarm, and not coming on-line or ready. Install some paper. If there is paper, ensure it is correctly installed. Ensure that the on-line button has been selected and there are sensors along the paper path; if these are defective, they will cause the printer not to print. Check the sensors along the paper path and replace if necessary.

Printer Port Connection at PC is Faulty: Try printing plain text via DOS (you can even use the print screen command on your keyboard while you are in the DOS environment). If still not printing, refer to your printer's owner manual and perform a self-test of your printer. If the printer prints OK, the fault is more than likely going to be with the printer cable or computer, not with the printer.

Self-Test is OK and PC Printer Port is OK: Faulty PCA interface / main logic board. Repair input/output circuit on PCA interface/main logic board. Extensive damage may require that the main logic board to be replaced.

3.5.3 Grinding Noises

Obstruction along Paper Path: There might be something caught along the paper path.

Gear Train Assembly worn out or Damaged: Inspect the gear train assembly. If worn out or damaged replace the gear train assembly.

Toner Cartridge is Faulty: Remove the toner cartridge and inspect its gearing and rotation replace if necessary.

Main Motor Faulty: Inspect the main motor's gearing and rotation and replace if it is found to be faulty. More often it is not damaged gearing attached to the main motor will not be available as a separate entity, meaning that the entire main motor assembly will require replacement.

Fuser Unit Assembly Faulty: As with faults with poor print quality and paper jamming, the fuser unit assembly will also cause grinding noises. Grinding noises coming from the fuser unit assembly is a sure sign of worse things to come. Replace the fuser unit assembly. It is always better to replace the entire assemblies as well as the meshing gears that it links up with

Cooling Fans: Grinding noises coming from the cooling fans tell you only one thing that it is about to stop. Seized fans will lead to the printer to overheat and cause possible extensive damage and hence costly repair bills. Replace the faulty fan immediately when this occurs

3.5.4 Paper Jamming / Misfeeding

Paper Quality Poor: Ensure that the paper you introduce into your printer meets with the standards of the printer manufacturer and replace if necessary. This is especially important when dealing with transparencies and labels

Obstruction along Paper Path: There might be something caught along the paper path.

Toner Cartridge is Faulty: Remove the toner cartridge and inspect its gearing and rotation replace if necessary.

Fuser Unit Assembly Faulty: As with faults with poor print quality and paper jamming, the fuser unit assembly is one of the most common faults regarding paper jamming. It is always better to replace the entire assemblies as well as the meshing gears that link up with it.

Sensors Faulty: Faulty sensors along the paper path will not provide the printer with the correct information about the position of paper in the printer and hence the printer will fall into a false paper jam mode. Inspect and test the sensors and replace if necessary.

Paper Output Delivery Assembly Faulty: This fault will prevent the printer from exiting the paper. Replace the worn out paper output delivery assembly and do not forget to replace the gear train assembly as well.

Paper Registration Assembly Faulty: This fault will prevent the paper from being picked up in the first instance. Replace the paper registration assembly and it is a good idea to replace the paper path sensors as well.

Clutch Drive Assembly Faulty: Causes internal paper jamming to occur. Replace both the entire clutch drive assembly and the gear train assembly.

Paper Delivery Assembly Faulty: Causes internal paper jamming to occur. Replace both the paper delivery assembly and the gear train assembly.

3.5.5 Fuser Malfunction

Fuser Unit Assembly Faulty: A Fuser Malfunction or Error 50 means just that, the Fuser Unit Assembly is faulty and requires replacement. An easy way to tell if your Fuser Unit Assembly is indeed faulty is to feel for any warmth emanating from the Fuser Unit Assembly. A cold Fuser Unit Assembly will generally inform you that something is wrong with the Fuser Unit Assembly. It is always be "****", to

replace the entire Fuser Unit Assembly along with the Meshing Gears that it links up with.

AC Power Supply Faulty: The AC Power Supply supplies power directly to the Fuser Unit Assembly therefore although you may have power circulating throughout the printer; the circuit responsible in providing voltage to the Fuser Unit Assembly has failed. Unfortunately, faulty AC Power Supplies also cause damage to the Fuser Unit Assembly. Sometimes both of these devices require replacement.

DC Controller Faulty: The DC Controller is really the heart and soul of control functions within a LaserJet Printer. The DC Controller is also the cause of Fuser malfunctions. Repair or replace the DC Controller.

3.6 Summary

In this unit, we have discussed about trouble shootings of dot matrix printer, which is more popular for its very low cost consumables being used. We have also described different trouble shooting of Ink jet printer, which is also popular for its quality. Finally, we have introduced trouble shootings of Laser printer, which are popular for its cost and quality.

3.7 Model Question:

Write a short note on the following.

1. Can you explain trouble shootings of Dot matrix printer?

2. Write about the trouble shootings of Ink jet printer.

3. Describe the trouble shootings of Laser printer.

UNIT- 4: WORKING WITH PLOTTERS

Structure:

4.0 Objectives

4.1 Introduction

4.2 Types of Plotters

4.2.1 Based on Pen

4.2.1.1 Single Pen Plotter

4.2.1.2 Multiple Pen Plotters

4.2.2 Based on its Position

4.2.2.1 Flatbed Plotter

4.2.2.2 Drum Plotter

4.2.3 Special Purpose Plotter

4.2.3.1 Electrostatic Plotter

4.2.3.2 Thermal/Laser/Pencil Plotters

4.3 Plotter Control Languages

4.4 Feature of Applicable Paper

4.4.1 Dimensional Criteria

4.4.2 Weight of Paper

4.4.3 Smoothness of Paper Surface

4.4.4 Anti-static Coating

4.4.5 Standard Paper

4.5 Pen Standard for the Plotter

4.6 Installation of Plotter

4.7 Summary

4.8 Model Question

4.0 OBJECTIVES

At end of this unit, you will be able to:

- Understand about the basics of different type of plotters.
- Explain the various kinds of plotters and its components.
- Distinguish in between flatbed plotters and drum type of plotters.
- Know the concept various types of interfacing devices used to connect plotters.
- Analyze statistical feature of suitability of plotters accordingly our purpose.
- Generate awareness about working process of different plotters.
- Identify the various features of paper handling plotters.

4.1 INTRODUCTION

Plotter is also an output device like printer. We can use plotters to get hardcopy of the information stored inside the computer. We know that most of the printers produce image on paper by printing a number of dots, whereas a plotter draws image on the paper by using a pen and by drawing with this pen on the surface of the paper.

For example, if you want to draw a line on the paper then on the printer this line will be drawn by printing a number of dots together to give an impression of a line. Whereas, on a plotter, the plotter will first place the pen at the starting point of the line and move the pen or the paper until the pen reaches endpoint of the line, just as you would draw it by hand. This process of drawing makes the plotter more useful in places where user requires graphical outputs consisting of lines, circles etc. The printer is more suited for graphic output such as photographs etc. and for textual output.

Different types of plotters are available in the market to perform different type of paper application. The standard size of the plotters are available, these sizes show the maximum of paper that can be used in them. The actual plotter size may be a little bigger then the paper size that can be used on these plotter.

Some common paper size and the name used for these sizes are given in the following table.

Paper Type	Size (in inches)
A (Letter size)	8 x11
B (Double or Tabloid size)	11x17
C (Broad Sheet size)	17x22
D	22x34
E	34x44

For business applications such as plotting of graph for the business presentation, etc, the A size plotters provides very good quality output. For architectural or engineering application, one would require a minimum D or E size plotter.

4.2 TYPES OF PLOTTERS

Plotters can be divided into different types based on the number of pen used to draw image, or it can be divided into different types based on the way pen and paper movement takes place, it can also be divided into different types based on the printing method used by the plotter. Cost and different capability of the plotter such as number of colors, maximum paper size, printing speed etc. very much depend on the type of the plotter.

4.2.1 Based on Pen

Because of the number of pen used in plotters, following two type of plotter are popular in the field of computer.

- Single Pen Plotters
- Multiple Pen Plotters

4.2.1.1 Single Pen Plotter

Single Pen Plotter is more popular because it is cheaper as compared to other plotter available in the market. We will a single pen plotter and it can plot multiple colors by making the plotter pause after drawing all the image of one color and allow you to substitute the pen with a new color pen. In a single pen plotter, we manually insert the pen into a pen holder.

Penholder holds the pen that drawn and collects different pens, as they are needed. Main pen holder acts on the instructions supplied by the computer. On some plotters, the pen is placed in a circular arrangement that is called a carousel. On others, the pen is placed in a straight line arrangement.

Once image of one color is drawn, the pen is put back at the pen-changing assembly and a new color pen is taken by the plotting mechanism.

Most of the drawing software that require plotters for output, such as the CAD software have inbuilt optimization method where they print all of the single color image first and then go for the second color pen and draw all the image with the second color, this save the plotter from frequently changing the pen. Some plotters in the market are available as single pen unit which can later on be upgraded to multiple pen models.

This is a good option if you want to start with single pen plotter and later on want to have the facility to use multiple pens.

4.2.1.2 Multiple Pen Plotters

Multiple pen plotter are better than the single pen plotters, the cost advantage of single pen plotters keeps it alive in the market. If you have a single pen plotter you can

plot multiple colors by making the plotter pause after drawing all the image of one color and allow you to substitute the pen with a new color pen.

This method is quite frustrating if you are required to change frequently the pen, for example, when printing small multicolor presentation on overhead transparencies. If most of the time your requirement is multicolor printing then it is better to go for a multi pen plotter, which will automatically change the pen, whenever required. In a single-pen plotter, you manually insert the pen into a pen holder. In a multiple-pen plotter, you load the pens into the pen-changing assembly and the main pen holder.

Most of the drawing software that require plotters for output, such as the CAD software have inbuilt optimization method where they print all of the one color image first and then go for the second color pen and draw all the image with the second color, this save the plotter from frequently changing the pen. Some plotters in the market are available in the market as single pen unit which can later on be upgraded to multiple pen models.

Smaller plotters have more requirements for the multiple pen then bigger plotters because most of the small plotters are used to draw business presentation which requires multiple colors, whereas most of bigger plotters are used to draw architectural and engineering drawing where single color output is more common.

Advantage of Multiple Pen Plotters

Multiple pen plotters have one big advantage over the single pen plotters that one can automate the complete plotting process using these plotters. Once all the required color pens are loaded into the plotters and the paper is provided, the user need to only give the command to plot the required image.

The plotter will complete the entire plot without any assistance from the user. Same process with a single pen plotter may require a number of pen change by the user, user will need to sit and keep watch waiting for the pen replacement.

Disadvantage of Multiple Pen Plotters

The advantage of the multiple pens is also a disadvantage, as one need to keep all the pen ready with ink filled and exposed nib. This makes the chance of pen tip drying a very common problem. If one uses an ink that does not dry fast enough then the image on the paper may get smeared when the pen makes multiple moves over the printed image. One needs an ink that must dry quickly enough on the media so that lines are not subsequently smeared by the pen tip, yet the ink should not dry on the tip when the pen is waiting to be used. To reduce this problem of ink drying on the pen tip, one can buy plotters which will have option to cap the pen when the pen is sitting in the pen-changing assembly. This cap will be automatically removed when the pen holder takes the pen for drawing.

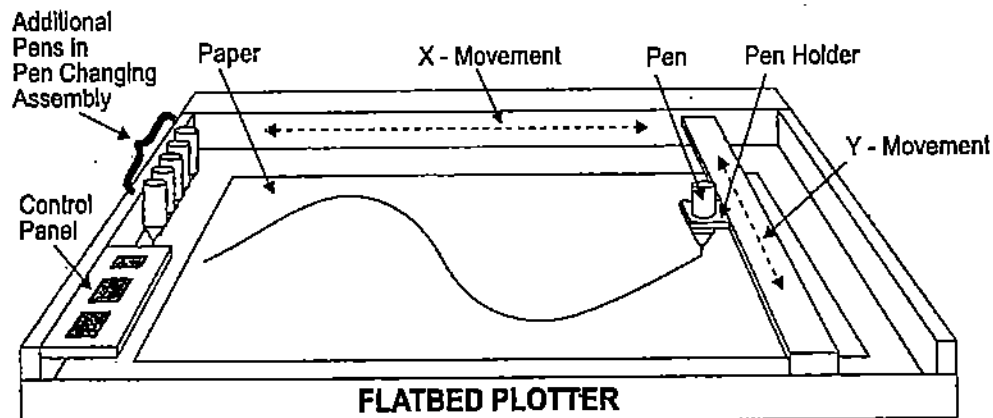
4.2.2 Based on its Position

Based on pen and paper interaction of plotter, the following two type of plotters are popular in the field.

- Flatbed Plotters
- Drum Plotters

4.2.2.1 Flatbed Plotter

On the flatbed plotter as the name suggests the paper lies on a flat bed i.e. a flat surface and the pen is moved on this paper to produce the required image. However, in the drum type plotter the paper is held on a drum like construction and both paper and pen are moved to plot the required image.



As shown in the above figure, flatbed plotters are flat rectangle shaped surface. These plotters are made bigger than the largest size paper that they can accommodate.

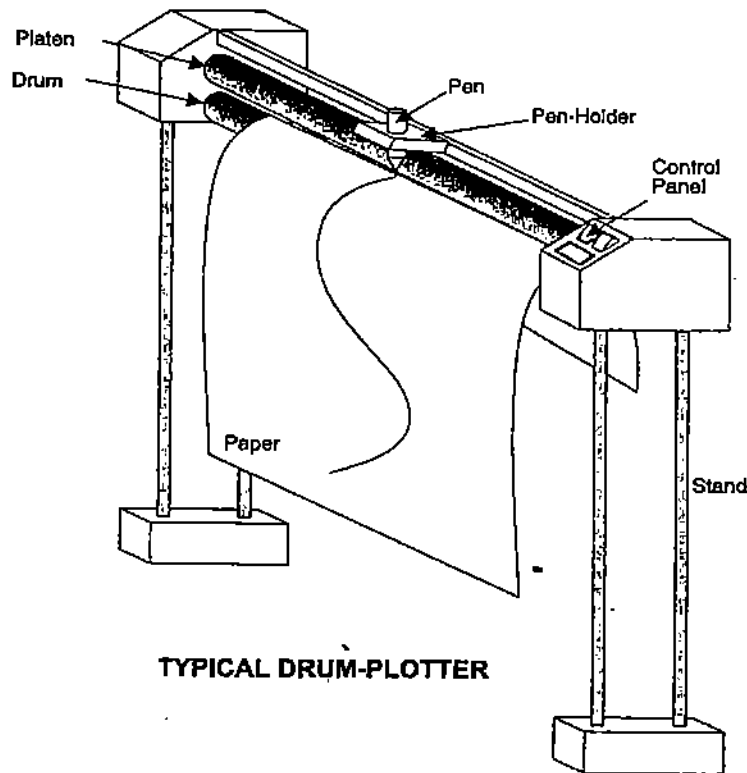
The extra space is required for plotting mechanism to move around. On this flat surface a pen is moved left, right, up and down to plot any required image. On a flatbed plotter the paper remains fixed on the plotter and the pen is moved.

As the paper does not move, one can produce image with very high accuracy on these plotters. This mechanical movement produces very high accuracy on the flatbed plotter's (often more accurate than 0.001 inch).

The fixing of the paper provided another advantage with these plotters i.e. one can plot image of maximum size limited to the plotter size but one can plot image of any minimum size required by the user, which is not possible on drum plotters, the drum plotters have fixed minimum size where it cannot plot any image smaller than that. These flatbed plotters are more suited in smaller of A and B size, on bigger size the space requirement to keep them makes it unsuitable, for higher size drum plotters are better option. The flatbed plotters plot the image as given in the figure. As shown in this figure, to plot any image the pen holder moves in the horizontal position from left to right and according to the image to be drawn the pen moves vertically up or down. The pen will touch the paper to draw the image or lift up from the paper where no drawing is required. This vertical and horizontal movement can be used to plot any image on the paper.

4.2.2.2 Drum Plotter

As shown in following figure, on a drum type plotter the paper is fixed between a drum and the flatbed, normally it is hold between two pinch rollers on both ends by rotating this platen and drum assembly one can move the paper up and down. When the paper is moving up and down the pen can be moved left and right horizontally to plot any shape on the paper.



4.2.3 Special Purpose Plotter

4.2.3.1 Electrostatic Plotter

The flatbed and drum type pen plotters plot the image by drawing on the paper. This produces a very good quality output, but the speed of these plotters are very slow for big drawings.

Normal pen speed of the plotters are 15 to 30 inches per second, this speed sounds very fast, but even a simple CAD drawing can contain thousands of line, and may require 1/2 to 1 hour to plot. One cannot increase the pen speed to make the plotter faster, as increasing of the speed will make the pen to skip areas on the paper, the high speed of the pen may make the pen to tear the paper. Electrostatic plotter is a good alternative to the pen plotters to plot drawing and engineering images. The speed of the electrostatic plotter is much more than the speed of the pen type plotters.

These plotters are basically bit image type of printers that can interpret the plot command received from the computer and produce the required image on paper.

5-4

After receiving pen movement instructions from the computer, these plotters create the image into their memory. Once the image to be plotted is made in the plotter's memory, the image is printed as a series of dots. Like a laser printer, the electrostatic plotters also have its own processor and memory to convert the instruction send by the computer into a proper image.

Normal speed of an electrostatic plotter is about .5 to 1 square inch per second. This speed sounds slower than the speed of the pen type plotters, but an electrostatic plotter will require same time without bothering about the number of lies to be plotted.

An electrostatic plotter can plot a D size image in 1 minute, which may require about 30 minutes on a normal pen plotter. Working principle of the electrostatic plotter is similar to the working of the laser printer, in these plotters a specially coated film or drafting film is used as printing media. On this printing media the image is drawn by using electrostatic charge by the plotters head.

Later this media is passed through toner. The toner is picked up by the charged area of the printing media and a visible image is made. This image is made permanent by fusing the toner on the printing media. As these types of plotters are very costly, they are used in areas where the number of drawing to be plotted is very high.

For higher throughput these plotter use paper rolls rather than the cut sheet used in the other type of plotters.

4.2.3.2 Thermal/Laser/Pencil Plotters

Thermal and Laser plotters are of two more types of plotters available in the market. These plotters work on the same principle as the electrostatic plotters.

These plotters generate the image to be plotted in their own memory as a dot matrix image and then print them using thermal or laser printing method. Thermal plotters require special paper on which the plotter generates the image by drawing dots with heat. On a laser plotter the image is generated on a drum by a laser beam, which is later transferred to the paper.

Pencil plotters are another plotter available in the market. These plotters are similar to the pen plotters with one difference that these plotters use pencil instead of pen to draw image on the paper. The advantage with these plotters is that one can erase any area of the image produced by these plotters for any correction.

4.3 Plotter Control Languages

Plotters move the pen or pencil according to the instruction received by the computer. These instruction are send in a language called as plotter control language. Plotter control languages include instructions such as to raise and lower the pen, move the pen from point A to point B and reset the plotter etc.

The most common plotter control language in the computer field is HPGL or Hewlett Packard Graphic Language. This HPGL is also implemented in many of the

Hewlett Packard's laser printers where the laser printer can act as a plotter and print the image send in HPGL language.

Almost all the CAD programs have facility to plot an image on plotters supporting HPGL language. Other than HPGL another DMPL (Digital Microprocessor Plotter Language) and HILOT (Houston Instruments Plotter Language) are used by the Houston Instrument in their plotters.

This DMPL has many added functions compared to the HPGL language for example the DMPL has built in fonts, single command for closed area fill etc. which is not available in the HPGL language. Whatever is the language support offered by a plotter, almost all the plotters have HPGL support, so if your CAD or other drawing software can plot using HPGL language then you can be sure that it will plot on any plotter available in the market?

4.4 Feature of Applicable Paper

We have seen that the accuracy of the plotter very much depends on the paper used for plotting of image, so one must choose the paper for plotter very carefully. The paper type that one must choose depends on the plotter being used, the pen being used and also on the job type. As the cost of this paper varies too much one can use cheap papers for the draft copy and use special papers when the final plot is taken out.

4.4.1 Dimensional Criteria

The dimensional stability refers to the expanding and contracting of the paper due to the environmental conditions. For example, the humidity or the temperature makes the paper to expand or contract.

The paper without proper dimensional stability may not produce proper image when used in a plotter. Also images that require to be drawn to the scale should be plotted on a paper with high a degree of dimensional stability.

4.4.2 Weight of Paper

A drum type of plotter needs to move the paper many times up and down during the plotting of the image, if the weight of the paper is very light i.e. if the paper is very thin then it may tear during the plotting process, also a thin paper may get torn by the pen nib of the plotter.

4.4.3 Smoothness of Paper Surface

Surface smoothness of the paper should be decided based on the plotter on which the paper is going to be used.

For a pen plotter one should use paper with smooth surface to produce clean image. Using of smooth paper also reduce the chances of pen tip wear. With a pencil plotter one needs a rougher paper surface, on smooth surface the lead used in these plotters will not be able to draw. If the paper surface is very smooth then the lead may just slide

away instead of drawing any image. Whatever to be the paper surface roughness, it should be constant across the entire surface to produce revert lines.

4.4.4 Anti-static Coating

Because of the frequent moving, the plotter paper can produce static electricity. This makes it necessary for a good quality plotter paper to have special anti-static coating on its surface.

Most of the plotters have some type of static eliminator to reduce this problem, but using, paper with anti-static coating further reduces any chance of the static electricity buildup.

Static electricity can damage the electronic components of the computer and the plotter and the plotted drawing will attract dust particles because of the static electricity on them.

4.4.5 Standard Paper

Some plotters such as electrostatic, thermal etc. require special paper to produce image. If you use ordinary papers on these plotters then you may not get any image plotted at all. Laser plotters require special papers that can withstand the high temperature of the fuser chamber used to bind image to the paper.

4.5 Pen Standard for the Plotter

With the paper, another very important part of the plotter is the pen. Many different types of pens are available for the plotters, from very cheap fiber tipped pen to very high quality, high cost jewel nib pen.

An inexpensive paper with fiber tipped pen can be used to plot the draft copies of the drawings and a good quality paper with jewel nib pen can be used to take out the final plot. One need to keep the pen capped most of the time to reduce the chance of ink getting dried and making the pen useless.

Most of the good plotters provide a self capping facility where a cap is placed on the pen, when the pen is not in use and this cap is automatically removed when the pen is taken out for plotting.

Using proper quality paper and pen is very important to get clean, clear image as output. It is very frustrating if you have to take another plot just because the previous output got torn or because of some smear of the ink.

4.6 Installation of Plotter

Most of the plotters use serial configuration this makes them very difficult to install and configure properly.

Each application used to plot an image may require a different plotter configuration because the serial port may be used in different ways by different software.

Recently many plotters provide option to connect them to the parallel port, which reduces many of the problems associated with the serial interface configuration.

For properly installing a plotter you must take care of the following points.

- List all the programs you want to use for the plotter.
- Check each programs installation and configuration information in detail for the plotter you want to use. (List the serial port communication parameter, cable wiring etc.)

Once you have enough information, use them to setup the plotter and the computer to follow a standard communication protocol.

One should use the cable supplied by the plotter manufacturer to connect the plotter to the computer, this should make the plotter work without any trouble. If you do not have proper cable, use the information provided in the plotter manual to make a cable.

4.7 Summary

In this unit, we have discussed about plotters, which is only responsible to produce output such as graph in the form of hardcopy. We have also described different type of plotters used in computer field. We have introduced working process of different plotters. Finally, we have explained about features of paper and for taking graph with plotters.

4.8 Model Question

Write a short note on the following.

1. What are roles of plotters? Can you explain flatbed type of plotters used in computer?

2. How multiple pen plotters are suitable? Explain its different advantage.

3. Describe the special purpose plotters and its installation.

NOTES

