Data Communications and Internet

Definition and Objectives

Computer-based communication refers to the movement of data and information generated by computers from one location to another via various types of communication lines. Computer-based communication systems often are called communication networks. The networks allow users to collect data and information from and transmit these to powerful central computers or terminals.

The *objectives* of communication networks are:

- To offer more timely interchange of information and data among users.
- To reduce the effort and cost required to collect and transmit business data and information.
- To support better performance of tasks and improved management control over an organization especially with several remote locations.

Why Study Communications?

Information society dominated by computers and communications is in front of us. In this society, the strategic resource is knowledge and information, and the value of that resource will be increased by the movement through the information pathway. A communication network is the pathway that allows information, data, and knowledge to be spread. Thus, the mainstream of the information age is communication.

Communications involve many technologies. In addition, many new communication technologies are being incorporated into the fabric of the information society as fast as people can learn how to maintain and use the technologies. This may be the most important reason for users to study communications.

Computer users who work in this society need to know technology incorporated into the communication networks. Although they use only applications involved in the communication networks, they still need to know how the communication networks work to improve their performance by using networks more efficiently and effectively.

GENERIC FORMS OF NETWORKS

Now let's see what are *generic forms of networks* that can exist in computer communications. A network can be defined simply as an interconnection of multiple teleprocessing devices. Three basic approaches to network processing exist in modern communications systems.

Timesharing

The earliest approach is known as *timesharing*. Timesharing occurs in networks in which multiple users gain access to a single host computer from their individual terminals. An example of timesharing is the use of a central computer at a nationwide company's headquarters by a variety of sales professionals dispersed around the nation.

Distributed Processing

The next approach gaining popularity in computer networking systems was the *distributed processing method*. This type of system involves the interconnection of minicomputers and microcomputers throughout a business

organization. Various types of resources are shared by the interconnected terminals, while the individual computers retain control over their own processing.

Client/Server Computing

The third approach to computer network processing, *client/server computing*, is currently attracting the most interest among management information systems professionals. The popularity of this approach is due to its ability to allow companies to selectively centralize and decentralize certain key operations. Thus, users can view client/server computing as a cross between the timesharing method (which emphasizes centralized computing) and distributed processing (which emphasizes decentralized computing). In client/server computing, the various applications available in the network are shared by several clients and one or more host computers (or servers). Clients obtain access to the network by means of desktop computers. The server, which can be a microcomputer, minicomputer, or mainframe, provides control for the entire network.

COMPONENTS OF A COMMUNICATION NETWORK

There are many components for a communication network. The basic three components are a host computer, a client, and a circuit. Besides these components, others include front-end processors, teleprocessing software, input terminals, and specialists.

Host Computer (Server)

The *host computer* or *server* (a mainframe, minicomputer, or microcomputer) is the main unit in the communication network. It stores data and programs that can be accessed by the clients. The host computer is responsible for performing the data processing tasks of the network. After processing has occurred, the data or information is routed back to a *front-end processor*. In distributed processing or client-server computing, several host computers may be tied together by the data communication network.

Client

The *client* is a hardware device (a terminal or microcomputer) at the opposite end of a host computer of a communication circuit. It usually provides end users with access to the data and software on the server. The client is an input/output device. The most widely used computer input/output devices for a communication network are the *terminals*. There are several types of terminals in common use in business today. The most familiar type of input/output terminal for most end users is the *keyboard terminal*. This terminal usually has a keyboard as an input device and a monitor and/or a printer as an output device. *Push- button telephones* can also be used as input/output terminals in computer-controlled communications networks. This type of terminal represents the most readily accessible and least expensive terminal in common use today. There are many special purpose terminals. One is a *point of sale (POS) terminal*. These terminals enable retailing institutions to instantaneously record sales data and update inventory records at the time of sale. Another is a *data collection terminal*. The term data collection terminal refers to a type of terminal which can collect information on the ongoing performance of factory employees. A *remote job entry* (RJE) *terminal* is another type. It is used in situations in which large amounts of information must be printed out at a remote location. A *cash register* used in many fast-food restaurants is another special type of terminal. It can update sales and inventory records.

Circuit

The *circuit* is a communication channel that allows the data and information to move through the pathway. It includes several types of lines, such as twisted wires, coaxial cables, and fiber optic cables. It also contains wireless transmission circuits such as microwaves and satellites, and it is becoming more popular.

Front-End Processor

The *front-end processor* handles the incoming and outgoing communications exchanges between the host computer and the peripheral terminals and other host computers both inside and outside the network. Thus, the front-end processor serves as both an input unit and output unit for the host computer. Usually, the front-end processor is a minicomputer, while a mainframe serves as the host.

Certain front-end processors can send messages from one terminal to another without involving the host computer. This ability is known as message switching. Front-end processors that perform message switching usually have a store and forward capability. When a terminal is unable to receive a transmission at a particular time, the front-end processor will hold the message in its secondary storage and resend it later.

Teleprocessing Software

Besides the hardware mentioned previously in this chapter, the host and the front-end processor must contain *software* in order for the network system to run effectively. Software may also be located in the terminals and the *cluster control units*. The types of software located in the host and the front-end processor will be discussed here.

- **TCM:** The teleprocessing software located in the host computer is known as the *TeleCommunications Monitor* (TCM). The typical functions of the TCM include: Setting priorities of incoming messages, providing network security, settling an interface between the teleprocessing system and the DBMS (database management system), and saving the contents of primary storage to secondary storage.
- **NCP:** The *Network Control Program* (NCP) is the name for the software located in the front-end processor. Its main functions are: determining if a terminal desires to obtain access to the network, organizing data and information within the network by stamping date and time, translating the codes received from other types of computer system, inspecting format errors in incoming delivery, saving recent incoming messages in secondary storage, and maintaining statistical records on network activity.

COMMUNICATIONS CHANNELS

Data communications can be carried from one location to another through a variety of communications *channels*. These communications media include telephone lines (twisted pairs), coaxial cable, fiber-optic cable, microwave, and satellite. In general, there can be two ways of connecting microcomputers with each other and with other equipment: guided and radiated media.
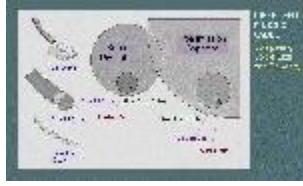
- **Guided Media:** Refer to channels that allow the transmission of data and information through a physical media such as a twisted pair wire, coaxial cable, or fiber optic cable.
- **Radiated media:** Refer to those that transmit data and information through the air such as microwave, or satellite.

The diameters and transmission capacities of the three kinds of cable are compared. See this Figure.

Twisted Pair

The telephone lines used to carry most of the voice and data communications consist of a pair of thin-diameter insulated copper wires (called twisted pairs). The wires are twisted around each other to minimize interference from other twisted pairs in the cable. Twisted pairs have fewer *bandwidth* than coaxial cable or

optical fiber. They have been the standard communication channels for voice, data and information, but are now diminishing because of more reliable media such as coaxial cable, optical fibers, microwave, or satellite.

Coaxial Cable

Coaxial cable can be used for telephone lines for transmission at a high frequency. Coaxial cable consists of a single core of solid copper. A coaxial cable can handle 80 times as many telephone transmissions as twisted pair media. Many computers in local area networks are linked by coaxial cables. Because of its sturdiness, coaxial cable is often used for telephone lines that must be carried under bodies of water. Because coaxial cables have very little distortion and are less prone to interference, they have low error rates.
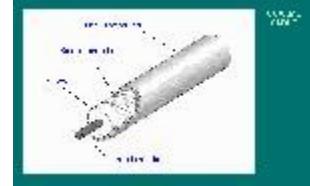

Fiber-Optic Cable

A fiber-optic cable consists of tubes of glass through which data are transmitted as pulses of light. Although a

fiber- optic cable is diametrically smaller than a human hair, it has 26,000 times the transmission capacity of twisted pair media. A major advantage of fiber-optic media is its high level of security. These communications channels are not susceptible to electronic interference. Therefore, they are a more reliable form of data transmission. Fiber-optic cables are also significantly less expensive than coaxial cable. A disadvantage of fiber-optic channels is that they cannot carry information over great distances.

Microwave

Microwave data transmission differs from the previously mentioned communications channels in that data is transmitted through the air instead of through cables or wires. Microwaves are high-frequency radio waves that can only be directed in straight lines. Consequently, microwave transmission is usually limited to communications occurring within the limits of a particular city or community. For microwave transmissions to be able to occur over larger distances, data messages must be relayed from one location to another using antennas placed at high altitudes usually twenty to thirty miles apart.
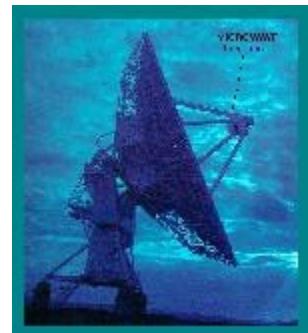
Satellites

Instead of antennas, satellites can also be used to transfer microwave messages from one location to another. Satellites rotate approximately 23,300 miles above the earth in precise locations. Satellite transmission stations that can both send and receive messages are known as earth stations. A major advantage of satellite transmission is that large volumes of data can be communicated at once. A particular drawback of several disadvantages is bad weather can severely affect the quality of satellite transmissions. Another one is that it has a serious security problem, because it is

DATA TRANSMISSION

Coding

In all digital communications channels, computers transmit data and information in forms of binary codes. Both sender and receiver of the data and information should have a standard for both to understand them.

A coding scheme for communications is a binary system, as in the computer systems. The system consists of groups of bits (0 or 1) that represent characters. In computer systems, a byte is a group of bits and represents a character. In data communications, a byte is the same, but some codes use different number of bits such as 5, 7, 8 or 9.

Two predominant coding schemes ASCII and EBCDIC. ASCII refers America Standard Code for Information Interchange. It is the most popular code for data communications and is the standard code on most communications terminals. Among two types of ASCII, a 7-bit code can make 128 character combinations, and an 8-bit can do 256 combinations. EBCDIC refers Extended Binary Coded Decimal Interchange Code. It is IBM's standard information code, and has 8 bits for a character.

Bandwidth

Each types of communications media has different transmission speed. The bandwidth is a measure of the transmission rate of communications channels.

- **Baseband:** Digital signals are commonly called *baseband signals*. Baseband is a communications technique in which *digital signals* are placed onto the transmission line without change in modulation. It transmits up to a couple of miles, and does not require the complex modems. Typical Token Ring and Ethernet use baseband signals.
- **Broadband:** Broadband is a technique for transmitting large amounts of data, voice and video over long distances simultaneously by modulating each signal onto a different frequency. Using the FDM (Frequency division multiplexing) technique, several streams of data can be transmitted simultaneously. Broadband is the bandwidth used for direct communication between very high-speed computers (e.g., large mainframe computers). This bandwidth includes microwave, satellite, coaxial cable, and fiber-optic media.

NETWORK TOPOLOGY

Networks can be classified by their topology, which is the basic geometric arrangement of the network. Different types of network configurations exist for network designers to choose from. Communications channels can be connected in different arrangements using several different topologies. This arrangement allows users to exchange information and share resources (software and hardware).

Four basic types of network configurations are star, bus, ring, hierarchical and mesh. Ring, bus, and star topologies are commonly used in LANs and BNs. Star and mesh topologies are commonly used in MANs and WANS. The networks are usually built using a combination of several different topologies.

- *Star*
- *Bus*
- *Ring*
- *Hierarchical*
- *Mesh*

Star Topology
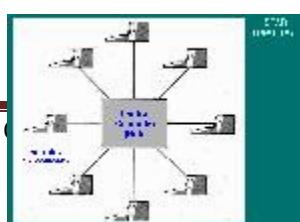
A star topology is one in which a central unit provides a link through which a group of smaller computers and devices is connected. In the star network, all interactions between different computers in the network travel through the host

computer. The central unit will poll each to decide whether a unit has a message to send. If so, the central computer will carry the message to the receiving computer.

Star networks represent a very popular form of configuration for time-sharing systems in which a central computer makes available resources and databases for several "client" computers to share. As such, the star network is appropriate for systems that demand centralized control. The disadvantage of the star network is that a processing problem in the central computer can be paralyzing to the entire system.

In a star network, the central unit may be a host computer or a file server. The h*ost computer* is a large centralized computer, usually a minicomputer or a mainframe. In contrast, the f*ile server* is a large-capacity hard-disk storage device. It stores data and programs files shared by the users on the network. Also, called a network server.

Bus Topology

In a bus configuration, each computer in the network is responsible for carrying out its own communications without the aid of a central unit. A common communications cable (the bus) connects all of the computers in the network. As data travels along the path of the cable, each unit performs a query to determine if it is the intended recipient of the message. The bus network is less expensive than the star configuration and is thus widely in use for systems that connect only a few microcomputers and systems that do not emphasize the sharing of common resources.
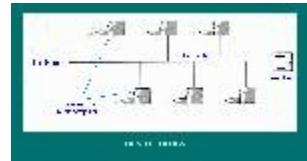
The problem in a computer on a bus topology does not frustrate the operation of the network, but a crack in the central cable will stop the whole network. Bus topology is popular because many computers can be connected to a single central cable. In a bus topology, each end user computer in the network handles its own communications control. There is no host computer or file server. As the information passes along the bus, it is examined by each terminal to see if the data is for it.

Ring Network
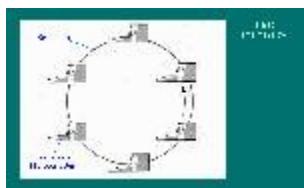
A ring configuration features a network in which each computer is connected to the next two other computers in a closed loop. Like the bus network, no single central computer exists in the ring configuration. Messages are simply transferred from one computer to the next until they arrive at their intended destinations. Each computer on the ring topology has a particular address. As the messages pass around the ring, the computers validate the address. If the message is not addressed to it, the node transmits the message to the next computer on the ring.

This type of network is commonly used in systems that connect widely dispersed mainframe computers. A ring network allows organizations to engage in distributed data processing system in which computers can share certain resources with other units while maintaining control over their own processing functions. However, a failure in any of the linked computers can greatly affect the entire network.

The ring arrangement is the least frequently used with microcomputers. However, as stated above, it often is used to link mainframes over wide geographical areas to build distributed data processing system. The loss of a mainframe usually does not restrain the operation of the network, but a cable problem will stop the network altogether.

Hierarchical Topology

A hierarchical network (or a tree network) resembles a star network in that several computers are connected to a central host computer (usually a mainframe). However, these "client" computers also serve as host computers to next level units. Thus, the hierarchical network can theoretically be compared to a standard organizational chart or a large corporation. Typically, the host computer at the top of the hierarchy is a mainframe computer. Lower levels in the hierarchy could consist of minicomputers and microcomputers. It should be noted that a system can sometimes have characteristics of more than one of the above topologies.

This topology is effective in a centralized corporation. For example, different divisions within a corporation may have individual microcomputers connected to divisional minicomputers. The minicomputers in turn may be connected to the corporation's mainframe, which contains data and programs.

Mesh Topology
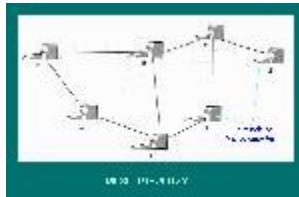
This is a net-like communications network in which there are at least two pathways to each node. In a *mesh topology*, computers are connected to each other by point-to-point circuits. In the topology, one or more computers usually become switching centers, interlinking computers with others.

Although a computer or cable is lost, if there are other possible routes through the network, the damage of one or several cables or computers may not have vital impact except the involved computers. However, if there are only few cables in the network, the loss of even one cable or device may damage the network seriously.

NETWORK TYPES

Communications networks differ in geographical size. Networks may be constructed within a building or across several buildings. Networks may also be citywide and even international, using both cable and air connections. There are three major network types: LAN (local area networks), MAN (metropolitan area networks), and WAN (wide area networks).

- *Local Area Networks*
- *Metropolitan Area Networks (MANs)/*
- *Wide Area Networks (WANs).*

Local Area Networks (LANs)

A computer communications network contained in a small area such as a commercial building is known as a local area network (LAN). A local area network's linkages usually are accomplished with either telephone, coaxial, or fiber-optic cables. Often, LANs use a bus configuration. The major benefit of a local area network is that it can help to reduce costs by allowing people and microcomputers to share expensive resources. Also, local area networks enable end users to participate in office automation technology. In addition, LANs may be connected to other networks by means of a network gateway. Network designers can choose the method in which messages are controlled in a local area network. Two basic methodologies are the token-passing approach and the contention-based approach. The token- passing approach allows the designers of a network to achieve a degree of centralized control. A group of data bits, or "token," is passed from one network node to another. A node can only send a message when it is in possession of the token.

In the contention-based approach, a node that wishes to send a message first listens to determine if another node is currently sending a message. If not, the node attempts to send its message. However, the lack of centralized control can result in a collision - two nodes attempting to send messages simultaneously. Consequently, contention-based approach is usually not suitable for networks with a large amount of communications activity.

Why Use an LAN?

There are two basic reasons for developing an LAN: information sharing and resource sharing.

- o **Information sharing:** This refers to having users who access the same data files, exchange information via electronic mail, or search the Internet for information. The main benefit of information sharing is improved decision making, which makes it generally more important than resource sharing.
- o **Resource sharing:** It refers to one computer sharing a hardware device (e.g., a printer) or a software package with other computers on the network. The main benefit of resource sharing is cost savings.

Types of LANs
By the categories, there are three common types of LANs. *Dedicated server LANs* account for more than *70 percent* of all installed LANs.

- *Dedicated Server Networks*
- *Peer-to-Peer Networks*
- *Zero-Slot LANs*

- **Dedicated Server Networks:** A *dedicated server LAN* can connect with almost any other network, can handle very large databases, have a dedicated network server, and uses sophisticated LAN software. Moreover, high-end dedicated server LANs can be easily interconnected to form enterprise-wide networks or, Sometimes, replace the host mainframe central computer. Generally speaking, the dedicated server is a powerful microcomputer. Three software components must work together and with the network hardware to enable communications: *the network operating system* (*NOS*) in the dedicated server, *the network communication software* that interconnects the server to the user computers, and *the application software* that runs on the server and client computers. Four common types of dedicated server LANs are file servers, database servers, print servers, and communication servers.
- **Peer-to-Peer Networks:** This network is a local area network that allows all users access to data on all workstations. In this networks, any computer can perform as both a client and a server. Each computer on the network shares its resource such as hard disk and printer with any other computer on the same network.
  This network is usually slower, has less capability, supports a limited number of computers, provides less sophisticated software, and is more difficult to manage than dedicated server LANS. However, this LAN uses cheaper computers and programs. Examples of peer-to-peer LANs include Artisoft's LANtastic, Novell's NetWare Lite, and Windows for Workgroups.
- **Zero-Slot LANs:** This LAN operates like peer-to-peer LAN, but offers limited, simple abilities such as sharing files and printers one another, transfer files, and transmit e- mail. It is inexpensive. It does not require a network interface circuit card. Its adapter plug can be plugged into a serial or parallel port. This network usually can handle up to 30 computers.

Metropolitan Area Networks (MANs) / Wide Area Networks (WANs)

MANs

The next larger network than LAN may be the *Metropolitan Area Networks* (*MANs*). MAN usually spans a geographical area that usually encompasses a city or county area. It interconnects various buildings or other facilities within this citywide area. For example, linkages can be established between two commercial buildings. A more recent use of MAN technology has been the rapid development of cellular phone systems.

WANs

A wide area network (WAN) is one that operates over a vast distance (e.g., nationwide). Its nodes may span cities, states, or national boundaries. This network interconnects computers, LANS, *BNs*, MANS, and other data transmission facilities. Typically, WAN will employ communications circuits such as long- distance telephone wires, microwaves and satellites. FOR example, nationwide automated teller machines used in banking represent a common application of a wide area network.

NETWORK APPLICATIONS

Computer-based communications systems allow end users to transmit data from one location to another via various types of communications lines. This type of communication has dramatically expanded the scope of business activities. The most vital component of any data communications system is the network. The network is any system that makes interaction between two or more computers possible. Today, computer users can tap into a vast number of resources to retrieve information on a variety of subjects and events. For example, a CEO of a nationwide corporation may wish to examine information about regional sales levels or expected economic conditions.

This section examines two types of network applications: *Groupware* and the *Information Superhighway* (*Internet*). These two applications are the future of information technology.

Groupware

This is a communications application that is growing in popularity (also called collaboration technology). Groupware is software that helps groups of people to work together more productively. Groupware allows people to exchange ideas, debate issues, make decisions, and write reports without actually having to meet face-to-face. The most important advantage of Groupware is its ability to help groups make decisions faster. This section focuses on four popular types of Groupware:

Electronic Mail

Electronic mail (*E-Mail*) is one of the earliest Groupware tools and is also the most heavily used tool today. E-mail is faster and cheaper than regular mail, and can substitute for telephone conversations in some cases. Several standards have been developed to ensure compatibility between different software packages. They include X.400, CMC (Common Messaging Calls), and MAPI (Massaging Application Program Interface).

Document-based Groupware (Notes)
Notes was the first document-based Groupware product. It is a document database designed to store and manage large collections of text and graphics to support ongoing discussions.

Group Support Systems (GSS)
Group Support System, one of the most popular uses of Groupware, is a software tool designed to improve group decision making in special purpose meeting rooms that provide networked computers and large screen video projection systems. These rooms are equipped with special-purpose GSS software that enables participants to communicate, propose ideas, analyze options, and evaluate alternatives. A group of employees

could view a document displayed on a projector screen. Then, each employee could manipulate the data from his station while others observed the resulting changes.

Video Teleconferencing

This is a video conference among several users, which is provided by one or more video cameras and several display monitors set up in special purpose meeting rooms. It provides real-time transmission of video and audio signals to help people in different locations to have a meeting. The advantage of this technology is the time and cost savings. Another form of video teleconferencing, desktop video conferencing, is growing up fast.

Information Superhighway (Internet)

The *information superhighway* or simply *Internet* is one of the most important developments in the history of information systems. It is growing fast. The Internet is not one network, but tens of thousands of networks linked together. In other words, it is a large network made up of thousands of smaller networks. Internet does not have overall central administration, because it is a collection of thousands of smaller networks. Internet

provides four basic functions to its users:
The other "back" button is at the end of this section.

- *E-mail*
- *Remote Login*
- *Discussion Groups*
- *Information Resources*

E-mail on the Internet

Electronic mail (E-mail) is one of the most rapidly growing developments in networked communications. Users of E-mail have their own file stored on a computer system. This file can be called a "mailbox." Access to a person's mailbox is protected by means of a password. Once logged on to an E-mail account, an end user may send messages and files to other mailboxes. An individual electronic mail transmission may be sent to one or many recipient accounts. This person may also read messages that have "arrived" in her or his mailbox from other E-mail accounts.

One of the main advantages of using E-mail is confidentiality. No one can access an individual's mailbox without knowledge of the password. E-mail also allows recipients to know the exact transmission times for each incoming message. Popular uses of E-mail have been to set up meetings within business organizations and to distribute memoranda throughout an organization. E-mail is increasing in popularity for communication between businesses. The main reasons for this are speed and cost; there is no reason to spend postage on a letter that will take three to seven days to arrive when electronic mail could be used for virtually nothing and arrive instantaneously.

Anyone with access to the Internet can send E-mail to anyone else on the Internet. Internet E-mail addresses have two parts, the individual user's account address and the address of the computer. The computer's address in turn has two parts, the computer name and its domain. The general format is therefore: *user@computer.domain*. Note that the "at" symbol (@) separates the user's account from the computer address, and that a period separates the name of the computer from its domain. Some computer names also have several parts separated by periods, so some addresses may have the format: *user@computer.computer.computer.domain* (e.g., user@sunset.backbone.olemiss.edu)

Remote Login (Telnet)

Internet offers a large number of services. Among them, the service that allows users to connect to a remote Internet host is called Telnet. Users on one computer in the Internet can login into other computers on the Internet by a special program, called Telnet, on your computer. This program uses the Internet to connect to the computer users specify. The users should know the account name and password of the remote computer. In Telnet, a user's computer is called the *local* computer. The other computer that the Telnet program connects is called the *remote* computer. An example of using Telnet is that users can read and send E-mail while traveling.

Discussion Groups

Discussion groups are lnternet users who have joined together to discuss some topic. There are many discussion groups on every topic imaginable, from cooking to biological science. Two groups are commonly used for business.

o **Usenet Newsgroups:** Usenet is a large collection of discussion groups involving millions of people from all over the world. This is the most formally organized among the discussion groups. To read Usenet articles, users use a program called a *newsreader*. There are a number of different newsreaders. In UNIX, **nn**, **rn**, **tin**, and **trn** are the most popular newsreaders.

o **Listserv:** A listserv is simply a mailing list developed on the large Bitnet network (not on the Internet). The listserv processor processes listserv commands such as requests to subscribe and unsubscribe, while the listserv mailer mails any message it receives to everyone on the mailing list.

Information Resources

The major use of the Internet is to find information. There are six major ways to find and achieve information:

FTP
Archie
Gopher
Veronica
World Wide Web

WAIS

- **FTP:** File Transfer Protocol (FTP) is the underlying set of specifications that support Internet file transfer. In other words, FTP is a service that allows us to copy a file from any Internet host to any other Internet host. A *ftp* program acts as a client and connects to the FTP server on a remote host. Internet users can download (copy files from the remote computer into users' computer) and upload (copy files from users' computer to a remote computer) files. An anonymous FTP site permits any Internet user to login using anonymous as the account name and a user's mailing address as the password.
- **Archie:** Throughout the Internet, there are a number of computers, called *Archie servers*, which provide a service to help users find the name of Anonymous FTP hosts that carry a particular file. Archie is a tool that allows users to search most of the publicly available anonymous FTP sites worldwide for specific files of interest.
- **Gopher:** The Gopher is a powerful system that allows users to access many resources of the Internet in a simple, consistent manner. To use the Gopher, all users need to do are making selections from a menu. In other words, Gopher is a menu-based tool that enables users to search for publicly available information posted on the Internet. The power of the Gopher lies in the fact that the resources listed in a menu may be anywhere on the Internet.

- **Veronica:** Veronica is a Gopher-based resource that users can use to search gopherspace for all the menu items that contain specified words. Veronica is to Gopher what Archie is to FTP. It enables users to search all publicly available Gopher sites by specifying key words.
- **World Wide Web:** WWW (World Wide Web) or simply Web is one type of information resource that is growing even faster than Internet itself is the World Wide Web. WWW is an attempt to organize all the information on the Internet as a set of hypermedia documents. Besides that, the WWW allows users to access all kinds of Internet resources, just by using a browser to read the appropriate document. The Web provides a graphical user interface and enables the display of rich graphical images, pictures, full motion video, and sound clips.

A *Web server* stores information in a series of text files called pages. These text files or pages use a structured language called HTML (Hypertext Markup Language) to store their information. HTML enables the author of a page to define different typestyles and sizes for the text, titles, and headings, and a variety of other formatting information. HTML also allows the author to define links to other pages that may be stored on the same Web server, or on any Web server anywhere on the Internet.

What Is the Internet?

> It seems like everyone's talking about the Internet these days. But what is it really? How does it work? How do you access it? And most important, what can it do for you at work or at home?
>
> Fortunately, accessing and using the Internet is fairly simple. Let this tutorial be your guide to the Internet as you connect for the first time and explore the network's vast and useful resources.

Introduction

This is a worldwide network that connects a vast chain of computer networks (private and public), which in turn link computers in homes, schools, businesses, and academic institutions throughout the world. It makes it possible to have all types of information & service for millions of people.

Historical Background

The Internet was the result of some visionary thinking by people in the early 1960s who saw great potential value in allowing computers to share information on research and development in scientific and military fields. J.C.R. Licklider of MIT, first proposed a global network of computers in 1962, and moved over to the Defense Advanced Research Projects Agency (DARPA) in late 1962 to head the work to develop it. Leonard Kleinrock of MIT and later UCLA developed the theory of packet switching, which was to form the basis of Internet connections. Lawrence Roberts of MIT connected a Massachusetts computer with a California computer in 1965 over dial-up telephone lines. It showed the feasibility of wide area networking, but also showed that the telephone line's circuit switching was inadequate. Kleinrock's packet switching theory was confirmed. Roberts moved over to DARPA in 1966 and developed his plan for ARPANET. These visionaries and many more left unnamed here are the real founders of the Internet.

Internet Technology

One of the greatest things about the Internet is that nobody really owns it. It is a global collection of networks, both big and small. These networks connect together in many different ways to form the single entity that we know as the Internet. In fact, the very name comes from this idea of interconnected networks.

Since it's beginning in 1969, the Internet has grown from four host computer systems to tens of millions. However, just because nobody owns the Internet, it doesn't mean it is not monitored and maintained in different ways. The Internet Society, a non-profit group established in 1992, oversees the formation of the policies and protocols that define how we use and interact with the Internet.

How it Works

Internet Protocol: IP Addresses

Every machine on the Internet has a unique identifying number, called an **IP Address**. The IP stands for **Internet Protocol**, which is the language, that computers use to communicate over the Internet. A protocol is the pre-defined way that someone who wants to use a service talks with that service. The "someone" could be a person, but more often it is a computer program like a Web browser.

A typical IP address looks like this: **216.27.61.137**

To make it easier for us humans to remember, IP addresses are normally expressed in decimal format as a *dotted decimal number* like the one above. But computers communicate in binary form. Look at the same IP address in binary:

11011000.00011011.00111101.10001001

How DNS works

Let's say that you type the URL www.howstuffworks.com into your browser. The browser contacts a DNS server to get the IP address. A DNS server would start its search for an IP address by contacting one of the **root DNS servers**. The root servers know the IP addresses for all of the DNS servers that handle the top-level domains (.COM, .NET, .ORG, etc.). Your DNS server would ask the root for www.howstuffworks.com, and the root would say, "I don't know the IP address for www.howstuffworks.com, but here's the IP address for the .COM DNS server."

Your name server then sends a query to the .COM DNS server asking it if it knows the IP address for www.howstuffworks.com. The DNS server for the COM domain knows the IP addresses for the name servers handling the www.howstuffworks.com domain, so it returns those.

Your name server then contacts the DNS server for www.howstuffworks.com and asks if it knows the IP address for www.howstuffworks.com. It actually does, so it returns the IP address to your DNS server, which returns it to the browser, which can then contact the server for www.howstuffworks.com to get a Web page.

One of the keys to making this work is redundancy. There are multiple DNS servers at every level, so that if one fails, there are others to handle the requests. The other key is caching. Once a DNS server resolves a request, it caches the IP address it receives. Once it has made a request to a root DNS server for any .COM domain, it knows the IP address for a DNS server handling the .COM domain, so it doesn't have to bug the root DNS servers again for that information. DNS servers can do this for every request, and this caching helps to keep things from bogging down.

Even though it is totally invisible, DNS servers handle billions of requests every day and they are essential to the Internet's smooth functioning. The fact that this distributed database works so well and so invisibly day in and day out is a testimony to the design.

1.3.3 Uniform Resource Locators

When you use the Web or send an e-mail message, you use a domain name to do it. For example, the **Uniform Resource Locator** (URL) "http://www.howstuffworks.com" contains the domain name **howstuffworks.com**. So does this e-mail address: **example@howstuffworks.com.** Every time you use a domain name, you use the Internet's DNS servers to translate the human-readable domain name into the machine-readable IP address.

Top-level domain names, also called first-level domain names, include .COM, .ORG, .NET, .EDU and .GOV. Within every top-level domain there is a huge list of second-level domains. For example, in the .COM first-level domain there is:

- HowStuffWorks
- Yahoo
- Microsoft

Every name in the .COM top-level domain must be unique. The left-most word, like www, is the host name. It specifies the name of a specific machine (with a specific IP address) in a domain. A given domain can, potentially, contain millions of host names as long as they are all unique within that domain.

DNS servers accept requests from programs and other name servers to convert domain names into IP addresses. When a request comes in, the DNS server can do one of four things with it:
- It can answer the request with an IP address because it already knows the IP address for the requested domain.
- It can contact another DNS server and try to find the IP address for the name requested. It may have to do this multiple times.
- It can say, "I don't know the IP address for the domain you requested, but here's the IP address for a DNS server that knows more than I do."

Definitions of key terms

- **Electronic mail (email**)- These are messages transmitted over the Internet or LAN
- **World wide web (www) -** part of the Internet that stores information in multimedia form - i.e. sound, photos, video and text.
- **Browser** -software that enables users to view web pages and move from one page to another
- **Bandwidth** - a measure of how much data can be sent through a communication channel in a specified span of time. It is measured in bits per second (bps)
- **Hypertext mark up language** (**HTML)** –a set of instructions used to program web documents and link websites
- **Universal resource locator** (**URL)** – a string of characters that points to a specific piece of information anywhere on the web. It is also the address of the web site
- **ISP-Internet service provider** –a company that connects you to the Internet
- **Website** – location of a web domain in a computer somewhere on the Internet
- **Web page** – document on the web including text, sound pictures and video. Websites contain web pages.
- **Home page -** is the site an Internet browser first visits when opened. A home page is often set to either an Internet portal or, if the computer is located inside an institution, the institution's web page.
- **Bookmark**
- **Spam/junkmail –** unsolicited mail **sent by some anonymous people**
- **Adware / Spyware**
- **Protocol**- communication rules that governs information exchange
- **Hits** – the number of times a subject/topic/keyword appears on the internet

Browsing the Web

Preparing to browse the web

The Internet is a massive repository of information spread allover the world. It is very easy to spend a lot of time and money browsing without getting good results. Therefore, before you embark on browsing it is very important to prepare. Your preparation will save you from wasting time looking at irrelevant WebPages. The following points should be very clear:

- The subject or topic to search – the kind of information you need from the internet
- The scope – avoid generalizations which will return irrelevant sites thus consuming more computer time
- Depth of detail -
- Expected sources of the information – websites, organizations, institutions etc
- Key search word – the is a term that is likely to return websites that best suit your subject or topic

In the absence of such preparation, you can "get lost" in an infinite web of "links of links"!

3.2 Using search engines:

3.2.1 Finding resources on the web

The Internet is rich in all kinds of resources. The challenge is finding the right website with the resources you require. Unless you know a particular website, **search engines** will be very useful. They are websites with software that can search for information anywhere on the Internet and retrieve it into your computer. The most common ones include:

- Google – www.google.co.ke
- Yahoo – www.yahoo.com
- Msn – www.msn.com
- Alta Vista – www.altavista.com

In the beginning

1990 *Archie*, developed at McGill University (Montreal) first search engine for finding and retrieving computer files. At the time these large institutional computers placed their data and program files into two categories: open and closed. When you 'logged-in' to another computer, you could access the 'open' files by identifying yourself as "anonymous" and using your e-mail address as the password. Then you could browse through their archive and download any files you wanted.

What **Archie** did was automatically at night (when the traffic was less) to visit all the archives they knew about and to copy the list into a searchable database (this piece of software was known as a *spider*). When you logged into an Archie site (by telnet) it would tell you where any file was and you could view and e-mail the results to yourself, and you could go through the entire 'log-in and retrieve' procedures for each computer yourself. It was a striking comment on the state-of-the-art at the time that McGill soon discovered that half of the whole US-Canada traffic was running through its Archie server, so it shut down public access. By then, however, there were many alternative sites hosting the service.

1991 The *Gopher* system represented an improvement on ftp retrieval developed at the University of Minnesota (whose mascot was a golden gopher).

The host computers (servers) put their files in a 'menu' form and the menus of the different servers were merged. Now you logged into any gopher server and you could query it for information by typing in keywords and, again like Archie, you would get a list of items. But now, instead of sending yourself the list and individually looking up the items, you scrolled down the list, pressed 'enter' and you were transferred directly to the relevant 'gopher' address, where you could read the contents. Then, if you wanted, you sent the file to yourself via e-mail. Since 'gopher' was a useful way for storing data, the system caught-on very rapidly. And within 'gopherspace' search-engine called *Veronica (supposedly Very Easy Rodent-Oriented Network Index to Computerised Archives)*, developed at the University of Nevada operated on the same principle as Archie but it also allowed you to distinguish between a search for 'directories' and an undifferentiated search combining directories and files (the latter was much larger and time-consuming). Again, having located something, you e-mailed it to yourself.1991 also saw the birth of *WAIS (Wide Area Information Server)* developed by Thinking Machines Corp.

Downloading Internet material:
Once you find the required information you may either read on the screen or make a copy on paper or storage media. Most websites avail their data freely, but some protect theirs from being copied without permission.

Presentation: web material is usually presented in any of the following formats:
- MS Word documents or any other text editor – any text editor can read this type of web material.
- Pdf – requires software called Adobe reader to read the material the way it is presented.
- Html – a mixture of text and web control characters.

Cut and paste

Ordinary text documents can be *lifted,* literally, from the WebPages onto a text editor like MS Word. Use the following steps:

- Using the mouse, highlight the portion of text that is relevant to you
- While the mouse arrow is pointing on the highlighted part of the WebPages, right-click the mouse and select: COPY
- Open the available text editor (e.g. MS Word) and open a new document
- Right-click inside the empty document and select paste
- The web page will be copied into the text editor. You can then manipulate like any other document

3.3.2 Saving WebPages

You can also save an entire web page. When this is done, everything that is on the screen will be saved. The effect is that several files and folders that accompany that page will be created so that when you open the saved file, it will appear just the way it was when it was viewed on the screen.

3.4 Marking favorite sites
Sometimes web material can be overwhelming. The search can return multiple options that all seem to be relevant. If you don't have enough time to go through all of them you can "bookmark" them under **Favorites** so that you can revisit them. Follow the following steps:
- While on the web page you wish to bookmark, click on the **Favorites** menu on the Internet browser
- Select the "**Add to Favorites**" option
- Select OK on the "dialogue box" the appears

3.5 Precautions – dos/don'ts, Best practices

Subscriptions – be careful

Other Internet Applications
E-commerce

This is a term that is used to describe conducting business – buying/selling- online using the Internet. The coming of the Internet has made it possible to conduct business from the comfort of your house, or cyber café. One is able to go shopping anywhere in the world and buy or sell anything without traveling. You will do business with faceless people or organizations, which you may never see or meet face to face.

This kind of business has brought with it some new concepts of doing business:

- **B2B-Commerce** – companies doing business online without Intermediaries. In most cases B2B involves large volumes of goods that are usually sold on wholesale basis in conventional modes of trade.
- **e-Money/e-Banking** – electronic movement of cash. No physical movement of money. All transactions (payments and receipts) are done electronically. Some banks offer mobile banking where clients can access their accounts and bills can be paid with instructions issued through their mobile phones.
- **e-TAILING** - conducting retail business online. There's more variety of Goods and services choose from. Buyers will benefit from product analysis software that enable them to tailor their products to their needs. This is definitely a better way of shopping than the conventional one where a shopper buys a product without knowing whether it will fit his specific needs.
- **Online auctions** – this includes the likes of eBay, which offer product from different people (merchants) for sale by biding.
- **Online adverts** – this is running advertisements on the Web. There are many marketing agencies offering this service. It is cheap and has a wider reach (i.e. worldwide) than traditional modes of marketing products.
- **Digital wallet** – software that imitates the physical wallet by holding the shoppers payment information. It also maintains a **digital certificate** the computer uses to identify the shopper and shipping information. The information in the wallet is encoded to protect the shoppers and the merchants against fraud. The wallet is usually installed on the shopper's computer.
- **Electronic billing purchase and payment (EBPP)** – a method of billing customers and collecting payments for the goods sold. There are two types: direct (which is done by a third party site) and consolidated – one-site collects all the shoppers bills and sends them to the shopper.

Requirements OF e-Commerce

Like any other business e-Commerce also has its requirements, which must be addressed before doing business can be possible. Among the primary requirements are:

- **Information technology infrastructure** – will include network-ready computer hardware equipped with relevant network software and Internet browser. In some cases internet-enabled or e-ready accounting software will be required to log in the orders and collect the sales revenues online.
- **Telecommunication** – without telecommunications, there would be no Internet and hence no e-Commerce. Telecommunication lines, modems, gateways/bridges, or switches will be required to facilitate connectivity to the Web.
- **Cashless banking** – with e-Commerce there is no physical movement or counting of cash involved. Everything is done electronically. In this regard the credit card reigns supreme. Those who have no credit cards find it really hard to conduct any real business on the Internet.
- **Courier services** – these are the companies, which facilitate the actual movement of physical goods that need to be delivered to clients in different parts of the world.
- **Human Resources** – these are the people who run the information systems on both side of the interconnection. The business selling must supply information concerning the available goods and services to the potential clients who will access it on the Web.
- **Software** – this is the driving force of e-commerce and there are quite a number of programs in the market for this business. Examples include: **product review software** (assists a shopper make his list by reviewing the products selected in terms of quality, quantity, price, etc), **shopping cart software**

(program that collects data on the items selected by a shopper and then submits them for billing), and **shopping search tools** (product dictionary offering categorized goods listing)

Present & Future of E-Commerce

The Internet has made the world to shrink into a small *virtual* village where everyone who is connected and everything that is online is within reach – literally, a mouse-click away! Initially the question was: do you have an email address? This has changed to what is your email address or what is your website/domain name.

By extension, E-Commerce that rides on Internet (cyber-highway) is posing a big threat to the traditional modes of doing business. The organizations which are preparing for e-commerce today stand a higher chance of success than those who are stuck in the old ways.

Local examples of e-commerce

- Virtual mobile phone credit card recharge offered by Standard bank
- Payment of bills
- E-banking – checking bank balances and payment of bills by phone as offered by the Co-op bank
- Power bill account Balance enquiries by KPLC

Electronic Mail (e-Mail)

- Launching the browser program
- o Mail addresses and mail box, Opening (signing in and accessing personal mail), closing mail box
- o **Email: Yahoo, MSN, Squirrel, GMail etc**
- Creating a new email address – signup
- Using WUCST mail service
- Using yahoo address book
- Forwarding, copying (cc, bc) mail to multiple addresses
- Security concerns: Taking care of Spam, junk mail, etc
- o Best practices

How Does the Internet Work?

The Internet is a worldwide collection of computer networks, cooperating with each other to exchange data using a common software standard. Through telephone wires and satellite links, Internet users can share information in a variety of forms. The size, scope and design of the Internet allows users to:

- connect easily through ordinary personal computers and local phone numbers;
- exchange electronic mail (E-mail) with friends and colleagues with accounts on the Internet;
- post information for others to access, and update it frequently;
- access multimedia information that includes sound, photographic images and even video; and
- access diverse perspectives from around the world.

An additional attribute of the Internet is that it lacks a central authority—in other words, there is no "Internet, Inc." that controls the Internet. Beyond the various governing boards that work to

establish policies and standards, the Internet is bound by few rules and answers to no single organization.

In February 1996, President Clinton signed into law the Communications Decency Act, which provides criminal penalties for those who post or transmit "indecent" material via the Internet. This law, however, has been challenged in U.S. courts by those who feel it would unfairly prohibit many legitimate uses of the Internet, and was ruled unconstitutional in July 1996. The federal government, however, is preparing an appeal.

The History of the Internet

Many people think that the Internet is a recent innovation, when in fact the essence of it has been around for over a quarter century. The Internet began as ARPAnet, a U.S. Department of Defense project to create a nationwide computer network that would continue to function even if a large portion of it were destroyed in a nuclear war or natural disaster.

During the next two decades, the network that evolved was used primarily by academic institutions, scientists and the government for research and communications. The appeal of the Internet to these bodies was obvious, as it allowed disparate institutions to connect to each others' computing systems and databases, as well as share data via E-mail.

The nature of the Internet changed abruptly in 1992, when the U.S. government began pulling out of network management, and commercial entities offered Internet access to the general public for the first time. This change in focus marked the beginning of the Internet's astonishing expansion.

According to a survey conducted by CommerceNet and Nielsen Media Research in early 1997, nearly one out of every four Americans over the age of 16 is an Internet user. And the number of users worldwide is believed to be well into the tens of millions. Other statistics are equally startling:

- A CNN report stated that Internet traffic in 1996 was 25 times what it was just two years earlier.
- The market research group IntelliQuest pegged the number of Internet users in the U.S. in late 1996 at 47 million - a 34 percent increase over the first quarter of that year.
- According to IBM, 146 countries currently have at least some level of Internet access.
- The technology research firm IDG estimates that by century's end, *one billion people* worldwide will have access to personal computers—more than doubling the computer-savvy population of 1996.

The Internet explosion coincides with the advent of increasingly powerful yet reasonably priced personal computers with easy-to-use graphical operating systems. The result has been an attraction of recent computer "converts" to the network, and new possibilities for exploiting a wealth of multimedia capabilities.

What Kinds of Information are Available?

In addition to text documents, the Internet makes available graphics files (digitized photographs and artwork), and even files that contain digitized sound and video. Through the Internet, you can download software, participate in interactive forums where users post and respond to public messages, and even join "chats," in which you and other users type (and, in some cases, speak) messages that are received by the chat participants instantly.

How Do People Use the Internet?

Obviously, the Internet can bring you a whole host of capabilities. But how can they be put to practical use?

Among the ways that users like yourself are taking advantage of the Internet are:

- Sharing research and business data among colleagues and like-minded individuals.
- Communicating with others and transmitting files via E-mail.
- Requesting and providing assistance with problems and questions.
- Marketing and publicizing products and services.
- Gathering valuable feedback and suggestions from customers and business partners.

The Internet's potential is limited only by users' vision and creativity. And as the Internet grows, new and innovative uses will surely follow.

The Sum of Many Parts

Unlike many computer networks, the Internet consists of not one but multiple data systems that were developed independently. The most popular and important systems are:

- **E-mail**, for exchange of electronic mail messages.
- **USENET newsgroups**, for posting and responding to public "bulletin board" messages.
- **File Transfer Protocol (FTP)**, a system for storing and retrieving data files on large computer systems.
- **Gopher**, a method of searching for various text-based Internet resources (largely obsolete).
- **TELNET**, a way of connecting directly to computer systems on the Internet
- **Internet Relay Chat (IRC)**, a system for sending public and private messages to other users in "real time"—that is, your message appears on the recipient's screen as soon as you type it.
- **CU-SeeMe**, a videoconferencing system that allows users to send and receive sound and pictures simultaneously over the Internet.
- **The World Wide Web**.

What Is the World Wide Web?

Of all the worthwhile features of the Internet, none has captured the public's imagination and contributed to the Net's growth so much as the World Wide Web. A method of posting and accessing interactive multimedia information, the "Web" is a true "information superhighway," allowing users the world over to access a wealth of information quickly and easily.

What Else Makes the Web Unique?

We have already cited hypertext and multimedia as two of the outstanding features of the World Wide Web. Among the Web's other distinct advantages are:

- **Open standards** that allow Web pages to be viewed through most any computer with an Internet connection and a Web browser.
- The ability to make software and data files available for **downloading**.

- **Ease of use**; just point and click on your browser to access the information you need.

For the individual or organization wishing to distribute information, the Web makes "publishing" easy and cost-effective. Putting information on the Web is quite inexpensive compared with traditional publishing, yet puts that information before a potential audience of millions. And unlike a printed publication, a Web document can be revised and updated at any time.

What Does the Web Empower Me to Do?

The World Wide Web brings you all of the advantages of the Internet—and more! Among the things that the Web allows you to do are:

- Publish and make frequent updates to on-line documents, and receive direct feedback on those documents via E-mail.
- Connect with a diverse, global audience—a community defined not by geography, but by interests.
- Keep abreast of the latest news and developments in your profession... often before such news appears in conventional media.
- Reference other Web resources easily via hypertext.
- Access otherwise obscure information not readily available in other media.
- Download useful software at little or no cost.

A Quick Tour of a Typical Web Site

A Web site may consist of a single Web page, or may comprise multiple, interconnected pages. The following is a view of how the CenterSpan home page appears through a popular browser:

1. **Browser tools** help you navigate around the Web. These vary among browsers, but most include tools for moving back and forth between pages, saving data to your hard disk, and controlling how Web pages will appear on your screen.
2. **A "Bookmark" list, "favorites" list or "hotlist"** lets you save the names and locations of favorite sites for easy reference (usually accessible through pull-down menus).

3. The **Uniform Resource Locator (URL)** is the address of a Web site. URLs for Web pages begin with the code http:// ("http" stands for hypertext transfer protocol).

4. Some Web pages contain special graphic **buttons** that, when clicked, take you to another resource as would a regular hotlink.

5. **Hypertext hotlinks** are connections to other pages and resources. To access, just click on them with your mouse or trackball. Most pages contain hotlinks throughout the text, as well as lists of links to other sites judged by the author/sponsor to be of particular interest. Some sites contain hotlinks to places where you can download software, listen to a sound file, or view a video clip.

6. CenterSpan uses several **applets**, or small program segments run by Java, ActiveX or some other protocol. Applets are downloaded when you access a Web site, and run only as long as you remain at that location. Applets can perform a variety of functions; the applet on the CenterSpan home page generates a scrolling banner, like the example below:

    If your browser does not have the capability to display Java applets, you will not be able to see this banner.

7. To make it easy to contact the author or sponsor of a Web site, most contain one or more **E-mail links**. Clicking on such a link brings up a form where you can type an E-mail message to a pre-set address (not shown here).

CenterSpan is among more elaborate Web sites that contain **forms**, where you can submit information in an easy-to-follow format, and **bulletin board** features that allow you to post information for others to see.

NB: THERE ARE HOWEVER SEVERAL DISADVANTAGES OF THE INTERNET

A] BRIEFLY EXPLAIN ANY FIVE OF SUCH DISADVANTAGES

B] INTERNET COMES WITH VIRUSES

- BRIEFLY DISTINGUISH BETWEEN VIRUS, WORM AND TROJAN HORSES

- EXPLAIN SOME OF THE CHARACTERISTICS OF VIRUS IN A COMPUTER SYSTEM